Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Liu P, Meng L, Normand EA, et al. Reanalysis of clinical exome sequencing data. N Engl J Med 2019;380:2478-80. DOI: 10.1056/NEJMc1812033

Supplementary Appendix

Table of Contents

Supplemental Methods

Clinical samples

Clinical cases were submitted for proband only diagnostic exome sequencing (ES). Parental samples, when available, were tested by targeted Sanger sequencing. Two cohorts correspond to two different 'data freeze' time points.[1,2] Cohort #1 consists of 250 consecutive exome cases received October 2011 to June 2012,[1] whereas Cohort #2 consists of 2000 consecutive cases referred June 2012 to November 2013.[2] The numbers of cases with both parental samples available for targeted testing are 199 and 1639, respectively.

Diagnostic criteria

We applied stringent clinical diagnostic criteria for establishing a molecular diagnosis. The criteria are in general consistent with those utilized previously[1,2] with accommodations implementing the published guidelines from the American College of Medical Genetics/ Association for Molecular Pathology (ACMG/AMP) regarding variant interpretation.[3] Specifically, all putative causative alleles are located in known 'disease genes' with well-defined clinical phenotype associations that importantly exhibit strong overlap with the patient's phenotype. For disease genes requiring one allele, i.e. monoallelic variants, the allele needs to be of a pathogenic or a likely pathogenic category based on the ACMG/AMP recommended guideline. For disease genes requiring two alleles, i.e. biallelic variants, at least one allele needs to be of a pathogenic or a likely pathogenic category and the other allele in trans can be of unknown significance or a higher diagnostic category. For cases

with complex phenotypes potentially contributed by more than one disease gene, if a diagnostic finding is found in only one gene, only those findings that explain the primary indication for clinical ES referral are considered as molecular diagnoses solving the case; the findings in genes explaining the patient's non-major complaints are considered as partial molecular diagnoses not molecularly solving the case.

Cases with a positive molecular diagnosis can be divided into three groups: 1) cases with a molecular diagnosis in the initial analysis and no more new diagnosis from reanalysis, 2) cases without a definitive diagnosis during the initial analysis but received one (or more) new diagnosis during reanalysis, 3) cases with an original molecular diagnosis but received more diagnoses during reanalysis. When the diagnostic yields are calculated, the number of unelated subjects/families are counted rather than the number of diagnoses. For example, a subject with two molecular diagnoses in two different genes is considered as one positive diagnosed case, rather than two. It should be noted that the diagnostic rates for two cohorts at the time the original exome reports were released are slightly lower than the numbers of diagnostic rates reported in the publications for those two cohorts,[1,2] because the calculations from the publications included some diagnoses made by reanalysis shortly after the initial analysis.

Manual reanalysis of exome data

Manual reanalysis was performed for Cohort #1 by American Board of Medical Genetics and Genomics (ABMGG)-certified clinical molecular geneticists following the complete standard review process used for new cases.[2] The process of reanalysis consists of variant reannotation and reinterpretation as well as assessing additional clinical information or blood samples/specimens from 'blood relatives' that were provided by the physician, when available, to implement the family based genomics approach. We also performed copy number variant (CNV) analysis using data existing from the exome pipeline (described below). Variant reannotation and reinterpretation were carried out using the Variant Call File (VCF) generated at the time of initial exome analysis. The variants were annotated using the knowledgebase at the time of December 2017. Distinctions between the new and the original annotations include gene-disease associations, inheritance patterns, literature regarding the gene or the variant, variant allele frequencies in various databases, and evidence accumulated in our in-house database regarding variant classifications. For selected cases, ordering physicians may provide updated clinical phenotypes or laboratory results for the patients after the initial analysis; ordering physicians may also provide DNA samples from additional family members to study the segregation of specific variants by Sanger sequencing. In such scenarios, the new clinical and segregation information will be considered in the context of the patient's original clinical description for variant-phenotype correlation; otherwise, variant-phenotype correlation will be performed using the original information provided. The standard procedure for evaluating variant classification remains the same as the standard used for the original exome reporting.[1,2] Candidate diagnostic findings were confirmed by Sanger sequencing and targeted familial studies were

performed using samples from family members submitted. Cases with a new molecular diagnosis were reassessed by an independent American Board of Medical Genetics and Genomics (ABMGG)-certified clinical laboratory geneticist and an American Board of Genetic Counseling (ABGC)-certified genetic counselor to confirm the agreement of the new diagnosis, and communicated with the ordering healthcare providers. This manual review process occurred systematically in December 2017 independently performing the reanalysis for all the individual 250 subjects case by case, and also sporadically during the five-year period after the initial report was released.

When a variant is found to be downgraded by reanalysis, the ES report is followed-up and communicated only when the variant downgrade overturns a previous diagnosis. Downgrades involving variants previously not considered to be the molecular diagnosis are not followed-up. For example, when a non-diagnostic variant of unknown significance is downgraded to a likely benign variant, an updated report is not issued unless there are other significant changes from the report.

Procedure of semi-automated reanalysis

For Cohort #2, in addition to the sporadic post-sign-out reanalysis activities, we have undertaken the reanalysis effort in December 2017 through a semi-automated approach. This pipeline consists of two parts, molecular filtering and phenotypic matching. To facilitate phenotypic-driven variant prioritization, we have implemented a phenotypic

match scoring system similar to our previously developed semantic similarity matching algorithm.[4,5] This system allows dynamic incorporation of new knowledge regarding gene-disease-phenotype mappings. Patient's clinical descriptions are summarized as structured languages and semantic similarity matching is performed for patient's phenotypes to the molecularly filtered variants. The molecular filtering takes a list of clinically annotated and prefiltered variants that is generated by the routine clinical review pipeline.[2] Then this set of variants is subject to more stringent population frequency filters to reduce the number of variants. The resultant filtered variants are subject to prioritization of a monoallelic genetic hypothesis and a biallelic genetic hypothesis. The prioritized variants are manually reviewed.

The complete R code for the semi-automated reanalysis pipeline embedded in an R markdown document is provided to the supplemental appendix. The markdown document provides detailed rationale and parameters for the workflow, step-by-step instructions as well as vignettes to run through the semi-automated reanalysis pipeline. The source code and mock patient VCF and HPO data are available on GitHub. The gene-based phenotypic matching and prioritization part of the code is developed into an online tool, PhenoMatcher. Web links for these resources are provided in the Data access section of this supplemental appendix.

Sanger confirmations and segregation studies were done similarly as reported in Cohort #1.[1] Cases with a new molecular diagnosis were assessed by two independent ABMGG-certified clinical laboratory geneticists and an ABGC-certified genetic counselor to independently confirm the agreement of the new molecular diagnosis, and communicated with the ordering healthcare providers.

CNV analysis

CNV analysis was carried out using two methods. For larger CNVs (>500 kb) we utilized data from cSNP analysis (Illumina HumanExome12 v1 array), that was performed as a quality control assessment to DNA fingerprint the sample for identification purposes, as part of the clinical exome pipeline. The cSNP array also allows a low-resolution genome-wide scan to detect CNVs. The cnvPartition 3.1.6 algorithm was used for CNV calling with the following parameters, confidence threshold: 35, minimum homozygous region size: 1Mb, minimum probe count: 3. Candidate CNVs were subsequently validated by comparing the exome read depth data in the same region with other control exomes. For smaller CNVs, we only focused on the detection of homozygous or hemizygous deletions based on complete absence of read coverage in exome data from a patient compared to normal controls. Homozygous deletion calls in known Mendelian disease genes related to the patient's phenotypes were selected as candidates. Candidate deletions were verified by an absence of PCR amplifications using specific primers located within the homozygous deletion.

Review time study for reanalysis

The review time for each case was considered to include all or a subset of the following segments: (1) reviewing a clinical summary that is synthesized from the clinical notes sent by the ordering physician, (2) performing *in silico* variant filtering and prioritization based on the semi-automated reanalysis procedure, (3) triaging candidate variants based on its phenotypic overlap with the patient and gross pathogenicity assessment (4) evaluating the pathogenicity of selected candidate variants that potentially fits the patient's phenotype, including reviewing the Sanger confirmation and family segregation data, (5) drafting clinical reports for cases with a positive finding, and (6) communicating positive results to the ordering provider. The length of time required for reviewing the clinical summary, drafting reports, and communicating results are estimated based on the re-review process of Cohort #1. Overall, the results of the cost analysis are presented as percentage of effort compared to an original exome analysis in the corresponding segment performed in the same diagnostic laboratory, rather than absolute numbers of variants or hours, which is not as meaningful as the relative portions of original effort because different diagnostic laboratories may have different sequencing protocols, interpretation guidelines, database contents, and reporting scopes.

The time investment for variant-level clinical correlation/triaging (step 3 mentioned above) is calculated based on the average number of variants prioritized by the semi-automated

pipeline as compared to the overall average number of variants subject to the original exome review from Cohort #1. The time investment for detailed reviewing of candidate variants (step 4 mentioned above) is calculated based on the number of Sanger sequencing reactions performed for reanalysis in the two cohorts. The rationale is that all the variants that are selected to be subject to Sanger sequencing confirmation and family segregation study are likely contributing related variants that require in-depth review of the disease mechanism, literature, and multiple databases.

When reanalysis does not yield any new reportable findings, time investment into report drafting and communication (step 5) is not needed. Therefore, we introduced a correction factor to reflect the percentage of cases that need to undergo these processes. These correction factors were calculated based on the positive rates of physician-initiated (percent of new positive cases from all the physician-initiated reanalysis requests, 33/290) or laboratory-initiated reanalysis (percent of new positive cases due to new disease gene or inheritance discovery and variant reclassification from all the cases in the two cohorts, 242/2250) generated from this study. With regards to the average time spent on report drafting and communicating the results, 30 reanalysis cases and 30 regular cases are timed for report drafting; 50 reanalysis cases and 50 regular cases are timed for communicating the results to the ordering care providers.

Clinical impact of exome reanalyses

The questionnaire listed below was sent to healthcare providers of 64 patients from Cohort #1 who received new molecular diagnoses from reanalysis. These 64 cases include 60 who were previously undiagnosed and 4 who had a previous molecular diagnosis but received additional diagnostic findings from reanalysis. Patient specific questions regarding the impact of reanalysis and difficulties in communicating the results collected from multiple providers for the same patient were consolidated into one data file. Among the patients for whom a response was not collected, efforts were made to determine whether the physician and genetic counselors were still at the same institution.

1. Did the updated exome report result in any of the following consequences in terms of patient management?

    - No change in medical management because patient is now deceased
    - No change in medical management
    - Additional diagnostic procedures completed
    - New medication started
    - Diet change instituted
    - Major procedure such as organ or stem cell transplantation completed
    - Palliative care initiated
    - Modifications to existing treatment/management
    - Relatives had genetic testing for known familial mutation(s) (KFM)
    - Information was used for reproductive planning such as testing of fetus by amniocentesis or CVS

- Other (e.g. results not communicated, please specify)

2. Were you able to communicate the updated results to the patient?

    - Yes
    - N/A because I was not involved in the communication process.
    - No (please specify the reason why results were not communicated, i.e. lost contact, unable to reach parents, did not attempt to contact, etc.)

3. How do you rate the level of difficulty in communicating the updated exome report to the patients who received a new diagnosis based on the updated report compared to communicating to patients regarding the initial exome report revealing a diagnosis?

    - More challenging in terms of establishing re-contact with the patient
    - More challenging because family does not seem to be aware that the updated report can be potentially important
    - More challenging in terms of explaining why the "negative" exome report issued before has become "positive" now
    - The level of difficulty is similar
    - Less challenging because the parents were prepared for possible updates to the results and knew what to expect
    - N/A because I was not involved in the communication process.
    - Other (please specify)

Healthcare providers' attitude towards exome reanalysis

A web-based survey was sent to the physicians and genetic counselors who were caring for the patients from Cohort #1 receiving new diagnosis by reanalysis. The questions in the survey are listed below.

A. Based on the notion that re-analysis can result in new diagnosis, would this affect the frequency that you schedule your patient's follow-up appointment after they had an initial "negative" exome report?

- Yes
- Depends on how often the laboratory allows/provides re-analysis
- No

B. Do you think that the clinical molecular laboratory should provide unsolicited updated exome reports that result in a new diagnosis (new diagnosis made usually by updated analysis pipeline, updated disease gene list, etc.)?

- Yes, the laboratory is obligated to do so.
- Yes, the laboratory can, but is not obligated to do this, as it will be difficult to define who should be responsible for the cost of this re-analysis activity.
- No, there is not a mechanism to communicate this information to the patient.
- No, other reason: (please specify)

C. If more than one provider is involved with the clinical exome analysis a patient, who should be notified when an updated exome report is available from re-analysis?

- All the providers.
- Only the provider who requested the re-analysis or who is most recently involved with the exome analysis
- None of the above: (please specify)

D. Do you think that the cost of performing routine exome re-analysis should be factored in the initial exome ordering?

- Yes
- No

E. Do you think that a fee should be charged when the referring physician request a re-analysis on an exome report?

- Always
- It depends on the frequency of the request
- Never

Supplemental Results

Manual and semi-automated reanalysis led to increased diagnostic rate

Systematic manual reanalysis of the Cohort 1 [1] increased the diagnostic rate from 24.8% in 2013 (or 22.8% excluding reanalysis activities prior to the 2013 publication time freeze) to 46.8% in December 2017, with 60 new cases being molecularly diagnosed (Figure 1). The contributing factors to the new diagnoses include a gene being newly associated with the clinical phenotype of interest after initial exome analysis (new 'disease gene', 75%), a variant being reclassified based on external or internal databases (variant reclassification, 5.9%), additional targeted variant segregation analyses on relatives of the proband (2.9%), updated clinical information provided by the referring physicians (8.8%), additional copy number variants (CNVs) (1.5%), and other reasons (5.9%) such as missed diagnosis due to failure to recognize or suspect the Mendelian condition based on patients' phenotypes relating to a gene.

Semi-automated reanalysis on the 2000 cases from Cohort #2[2] referred for ES subsequent to the initial Cohort #1 increased the diagnostic rate from 25.2% to 36.7% (an additional 230 cases, Figure 1). New molecular diagnoses resulted from new disease genes (64%), variant reclassification (14%), CNVs (6.5%), additional familial studies (6.5%), updated clinical information (3.6%), and other reasons (5.6%, including 4 cases with previously false negative calls due to inadequate NGS coverage) (Figure 1).

Calculation of clinical sensitivity for the semi-automated analysis approach

Since systematic manual reanalysis has been conducted on Cohort #1, this cohort is considered a 'gold standard' for clinically reportable variants. When calculating the clinical sensitivity, the total number of genes containing variants contributing to the overall clinical diagnoses is counted, rather than the number of patients (one patient may have more than one diagnosis) or the number of variants (diagnosis from recessive genes often are associated with two variants). Pathogenic variants contributing to secondary diagnoses not related to the patients' referral indications are not counted. Pathogenic variants contributing to a partial diagnosis of the patients' referral indications are counted, even though we do not consider these patients to be "molecularly diagnosed". Diagnoses contributed by large copy number or structural variations are not counted as they are handled through a separate mechanism as described in the following section.

There are 128 molecular diagnoses from Cohort #1, including 2 partial diagnoses and 1 diagnosis from CNV. Among the total of 127 molecular SNV/INDEL diagnoses from Cohort #1 (with the 1 CNV diagnosis excluded), 118 were identified by the semi-automated reanalysis procedure, resulting in a diagnostic sensitivity of 92.9%. The molecular details for the 127 diagnoses (including 154 variants) are provided in Table S1. Information regarding whether each variant is detected by the semi-automated reanalysis pipeline and potential reasons for a variant not passing through the pipeline filter is also provided in the table.

Medical knowledge accrual contributes the most to new molecular diagnoses

In both cohorts, the vast majority of new molecular diagnoses resulted from newly discovered disease genes (75% and 64%, respectively), consistent with the rapid pace of new disease gene discovery in the past few years.[6,7] New molecular diagnoses also resulted from upgraded variant-level classifications (n=38) in known disease genes, representing 5.9% and 14% of increments in diagnostic yield in Cohort #1 and Cohort #2, respectively (Figure 1 and Table S2). The upgraded variants were included in the original report as variants of unknown significance for 23/38 cases. Reclassifications were prompted by new knowledge of identical or allelic variants from literature reports or internal/external clinical diagnostic databases; one variant was upgraded due to recently reported expanded phenotypic associations (i.e. phenotypic expansion[8]) for a known disease gene (*FOXP1*).

Variants not included in the original report, but considered diagnostic after new evidence emerged (15/38), may present a major leap forward in diagnostics from a physician's and a family's perspective. The omission of these variants from the original report was complicated by several factors including variant-specific atypical phenotypic presentations (8/15), gene-specific multiple disease inheritance patterns and mechanisms (2/15), newly discovered isoforms encompassing previously unknown exons (2/15), and complex patient phenotypes obscured by multilocus molecular diagnoses (4/15).[5,8] These findings illustrate the utility of performing variant-level reclassification to facilitate the clinical detection of challenging clinical diagnoses.

New molecular diagnoses achieved by clinician-initiated efforts

Reanalysis is not always initiated by the genomics/molecular diagnostic laboratory. Clinicians may request that the laboratory reanalyze the exome data based on an evolving clinical picture. In the two cohorts studied here, 8.8% and 3.6% (combined n=14) of the new diagnoses were triggered by new clinical information provided by the clinician. Key information included biochemical data pinpointing a specific pathway or target gene (n=7) and more specific clinical phenotypes that emerged with age (n=7) (Table S3). The success rate of reanalysis requested by clinicians is 14% (21/154) when new clinical information was provided or 8.8% (12/136) when there was no update in the patient's clinical picture.

Submission of samples from affected or unaffected family members also aided with clarifying the pathogenicity of variants, contributing to 2.9% and 6.5% of all reanalysis-based new diagnoses from the two cohorts. Familial testing revealed *de novo* origin (n=10) or compound heterozygosity (n=7) to contribute to the new monoallelic (dominant disease trait) or biallelic variant (recessive disease trait) molecular diagnoses, respectively. An additional case (n=1) carrying a hemizygous change received an upgraded diagnosis after family segregation studies with multiple family members. In one case, the causative variant was transmitted from a symptomatic father, in whom the monoallelic variant was *de novo*. When additional family member samples are submitted to test for targeted known familial variants, the success rate of making new molecular diagnoses is 12% (18/146).

CNV analysis contributing to new diagnoses

As genomics and ES are quickly adapting to different translational medicine clinical scenarios, evaluation of CNVs using the ES raw data may produce unexpected molecular diagnoses. We performed retrospective CNV analysis for the two cohorts using the cSNP array data intended as a quality control step in the clinical exome pipeline. Diagnostic CNVs not previously known to the referral physician were identified in 1 and 19 patients from the two cohorts (Table S4). In addition, CNVs known to confer a disease risk with reduced penetrance, which do not qualify as diagnoses by current clinical genomics molecular diagnostic practice standards, but can still affect the patients' management, were identified in one and five patients from the two cohorts (Table S4). These CNVs were previously unidentified because earlier analyses of exome cases mainly focused on SNVs and did not interrogate potential causal CNVs.

In addition to large CNVs, the reanalysis also included detection of exonic level homozygous/hemizygous deletions based on the exome data using a recently developed bioinformatics approach, similar to the method described in Gambin *et al.*[9] Five exon-level homozygous deletions were identified, ranging from one to four exons in size (Table S4). This provided definitive molecular diagnoses and recurrence risk information for the family by demonstrating biallelic variants consistent with recessive inheritance and the Mendelian expectations of heterozygous carrier parents.

Reanalysis augmented diagnostic resolution by revealing more cases with multilocus genomic variations

Increasing use of genome-wide technologies such as ES has revealed that occasionally a patient's disease phenotype may represent a blended phenotype caused by multilocus pathogenic variation, leading to multiple conditions in one patient, with clinical features that may be *distinct* from each other or having *overlapping* features.[1,2,5,8] In the two cohorts, 23 additional patients exhibited dual or triple molecular diagnoses after being subjected to reanalysis (Table S5). In 22 cases, the original ES report had a single diagnosis, with reanalysis adding one (n=21) or two (n=1) previously unrecognized diagnoses. One patient has two new diagnoses identified consecutively after an initial negative exome report. More than half (13/23) of the new diagnoses present with overlapping features to the original diagnoses (Table S5); without comprehensive re-analysis, such new diagnoses may remain unrecognized, and not clinically suspected. In addition, seven patients received a partial molecular diagnosis from reanalysis, which suggests the potential existence of a yet-to-be-uncovered second molecular diagnostic finding. The percentages of multilocus diagnoses among diagnosed cases before and after reanalysis are 5.4% (3/56) versus 6.8% (8/117) for Cohort #1 and 4.4% (22/504) versus 5.4% (40/734) for Cohort #2.

Previous molecular diagnoses overturned by reanalysis

In the course of this ES reanalysis effort, we identified six previous molecular diagnoses that were overturned because of updated knowledge of population variant allele frequency inconsistent with rare disease (Table S6). One variant thought to contribute to

a previous molecular diagnosis was downgraded after familial segregation studies. These variants have been communicated to the healthcare providers with the updated clinical report.

Challenges in communicating results of updated diagnoses

The response rate from the healthcare providers regarding the outcomes of return of updated results was 66% (42/64) (Table S7). The time intervals between the updated report being issued and the time of this study ranged from 1 to 64 months; for 12 cases, the time interval was below six months. Updated results were successfully communicated to 30 patients in a follow-up appointment. For the remaining 12 patients, the results had not reached the patients or the patients were informed but did not arrive at their scheduled follow-up clinic appointment. Return of reanalysis-generated updated molecular diagnostic results was rated by genetic counselors as less challenging in 21 (70%) cases, more challenging in 7 (23%) cases and of similar difficulty in 2 (6.7%) cases when compared with communicating the original report.

Five years after ordering the initial exome analysis, 14% (9/64) of the original ordering physicians no longer practice at the same institution; an even higher percentage, 63% (29/46) of the genetic counselors have relocated. The personnel turnover in the medical genetics team, especially for the physicians, may create additional barriers in the communication process, as the absence of the original care provider may increase the risk of communication breakdown.

When more than one physician from the same or different disciplines are involved with utilizing the same set of exome data for patient care, it becomes a conundrum to whom the laboratory should direct the updated HIPAA-protected exome report. In this study, 47% (135/290) of all physician-initiated reanalysis requests were not from the same physician who ordered the original exome analysis.

Follow up of patients by physicians after receiving the updated clinical reports

Among 30 patients receiving a new diagnosis from reanalysis at a follow-up clinic appointment (Table S7), the clinical management plan was impacted for 17 patients, including new medication started (n=4), diet change instituted (n=1), redirection of goals of care requested by parents (n=1), modifications to existing treatment/management (n=10), and additional diagnostic procedures completed (n=8). In addition, reanalysis results triggered seven families to initiate genetic testing for known familial mutations in relatives and three families to use the new diagnosis for reproductive planning such as genetic testing by amniocentesis or chorionic villus sampling. In the remaining 13 patients, clinical management was not immediately affected. Nevertheless, this may change as the biological perturbations responsible for the phenotype emerge, and corresponding practice guidelines are formalized.

Survey results regarding healthcare providers' attitudes towards exome reanalysis

The survey was sent to 55 healthcare providers. The response rate was 42% (23/55). The survey results are listed below.

A. Based on the notion that re-analysis can result in new diagnosis, would this affect the frequency that you schedule your patient's follow-up appointment after they had an initial "negative" exome report?

- (6/22) Yes
- (9/22) Depends on how often the laboratory allows/provides re-analysis
- (7/22) No

B. Do you think that the clinical molecular laboratory should provide unsolicited updated exome reports that result in a new diagnosis (new diagnosis made usually by updated analysis pipeline, updated disease gene list, etc.)?

- (12/21) Yes, the laboratory is obligated to do so.
- (9/21) Yes, the laboratory can, but is not obligated to do this, as it will be difficult to define who should be responsible for the cost of this re-analysis activity.
- (0/21) No, there is not a mechanism to communicate this information to the patient.
- (0/21) No, other reason: (please specify)

C. If more than one provider is involved with the clinical exome analysis a patient, who should be notified when an updated exome report is available from re-analysis?

- (10/23) All the providers.
- (13/23) Only the provider who requested the re-analysis or who is most recently involved with the exome analysis

- (0/23) None of the above: (please specify)

D. Do you think that the cost of performing routine exome re-analysis should be factored in the initial exome ordering?

- (11/22) Yes
- (11/22) No

E. Do you think that a fee should be charged when the referring physician request a re-analysis on an exome report?

- (2/23) Always
- (14/23) It depends on the frequency of the request
- (7/23) Never

Cost of laboratory-initiated reanalysis

We analyzed the cost of laboratory-initiated reanalysis via the semi-automated reanalysis pipeline (Table S8), using an annual systematic reanalysis schedule[10] over five years. This results in five iterations of variant filtering and prioritization using updated curation databases. Based on the pipeline used in this study, ~1.8% of the total variants is estimated to be manually reviewed during each reanalysis, which translates to ~8.5% of the original review effort for sustained variant triaging over a five-year period. The in-depth re-evaluation of selected likely contributing variants is estimated to take a cumulative time period equivalent to ~13% of that spent for a corresponding task in the initial review. The

subsequent report drafting and communication steps are estimated to occur only for 242/2250 or 10.8% of all cases over five years (negative reanalysis results do not receive a full updated report). The positive cases do require more time investment for report drafting-estimated at ~250 % of time on average compared to that of a new exome report, considering the potential intricacies such as investigating the legitimacy of a new disease gene. Our time study showed no significant difference in the time spent for result communication between reanalysis and regular cases.

The cost components of a physician-initiated reanalysis are similar to that of an original analysis except for the sequencing cost and the report drafting time (Table S8). It is estimated that approximately 33/290 or 11.4% of all the physician-initiated reanalysis requests result in reports needing Report amendments and a report addendum.

Data Access

All variant data discussed in this study have been deposited in ClinVar.

GitHub link: https://github.com/liu-lab/exome_reanalysis

PhenoMatcher link: http://genomicanalysis.research.bcm.edu:3838/PhenoMatcher/

Supplemental References

1. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med 2013;369:1502-11.

2. Yang Y, Muzny DM, Xia F, et al. Molecular findings among patients referred for clinical whole-exome sequencing. Jama 2014;312:1870-9.

3. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405-24.

4. James RA, Campbell IM, Chen ES, et al. A visual and curatorial approach to clinical variant prioritization and disease gene discovery in genome-wide diagnostics. Genome Med 2016;8:13.

5. Posey JE, Harel T, Liu P, et al. Resolution of Disease Phenotypes Resulting from Multilocus Genomic Variation. N Engl J Med 2017;376:21-31.

6. Boycott KM, Rath A, Chong JX, et al. International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases. Am J Hum Genet 2017;100:695-705.

7. Posey J, O'Donnell-Luria A, Chong JX, et al. Insights into genetics, human biology and disease gleaned from family based genomic studies. Genet Med 2019;In print.

8. Karaca E, Posey JE, Coban Akdemir Z, et al. Phenotypic expansion illuminates multilocus pathogenic variation. Genet Med 2018;10.1038/gim.2018.33.

9. Gambin T, Akdemir ZC, Yuan B, et al. Homozygous and hemizygous CNV detection from exome sequencing data in a Mendelian disease cohort. Nucleic Acids Res 2017;45:1633-48.

10. Rehm HL, Bale SJ, Bayrak-Toydemir P, et al. ACMG clinical laboratory standards for next-generation sequencing. Genet Med 2013;15:733-47.

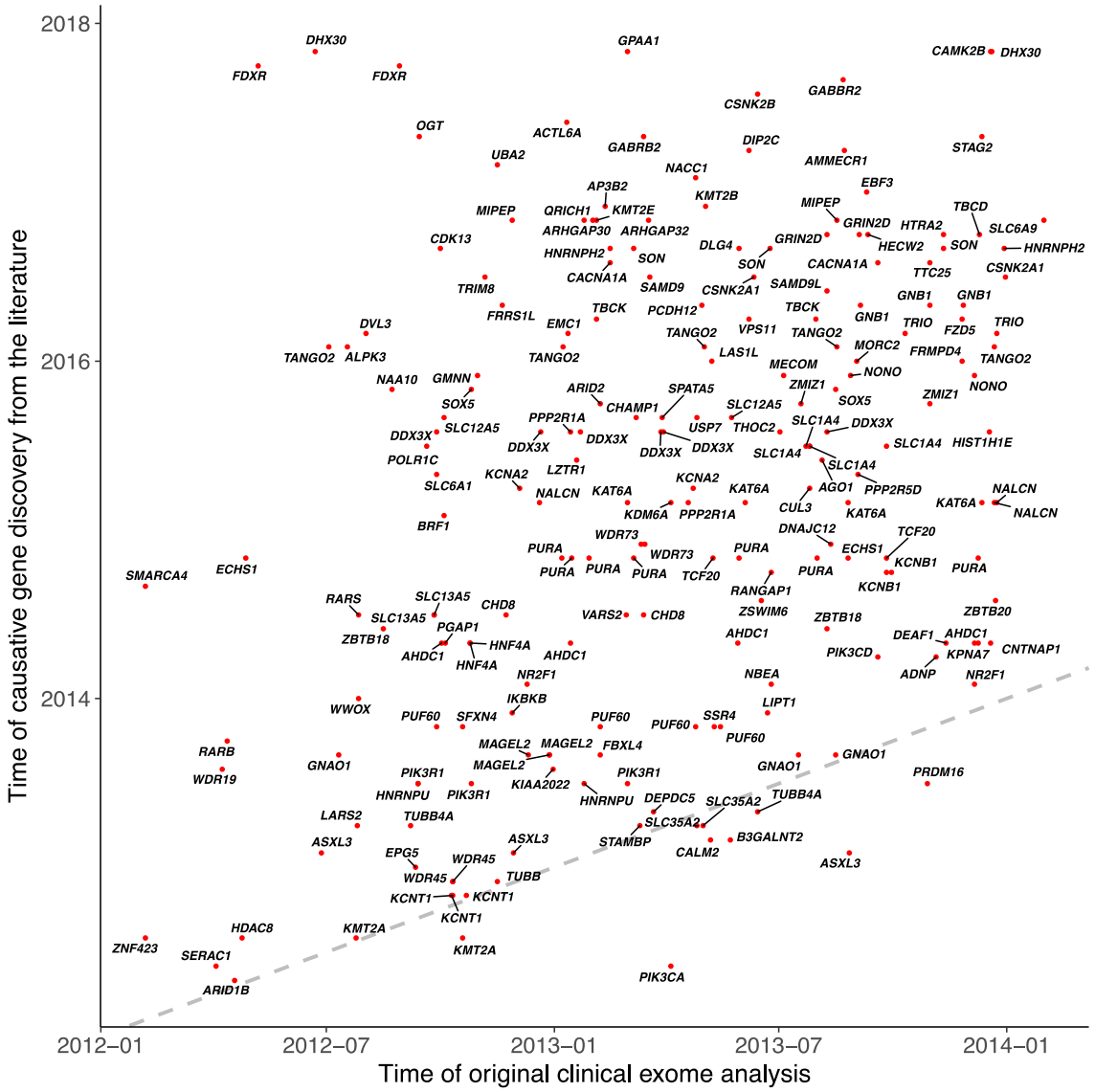

Figure S1. New molecular diagnoses contributed by newly discovered disease genes. Each dot represents a new molecular diagnosis from the two cohorts, with the causative gene name as labels. The X-axis denotes the time when the original exome report was released with non-diagnostic findings. The Y-axis denotes the time that the causal gene for the patient is linked to disease in the literature. If the red dot is located above the dashed line, the disease association has not been established at the time the initial clinical exome report was made. A few dots are located below the dashed line. This is a result of

the delay of the new disease association being added to databases such as OMIM and HGMD. Genes whose support for disease association solely derive from our internal database are not plotted.

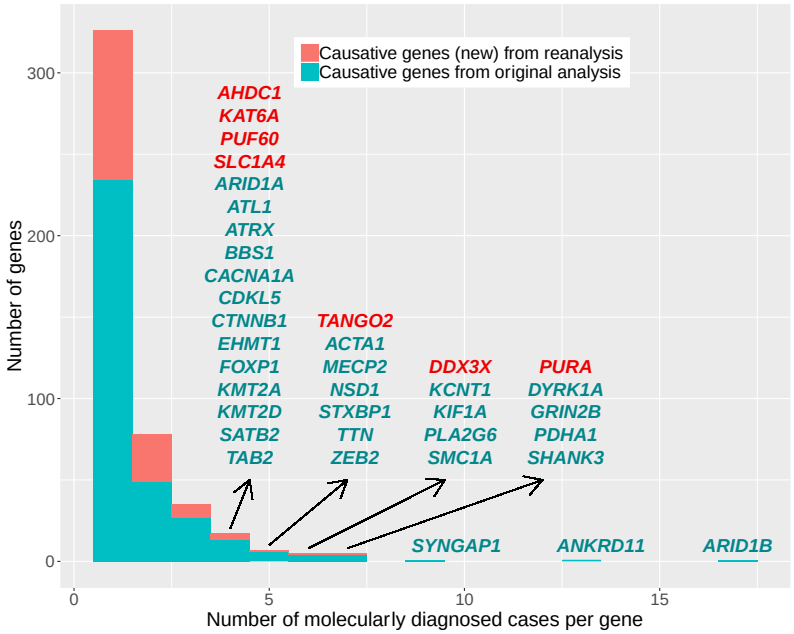

Figure S2. Mutation burden of genes contributing to diagnoses from the original analysis or the reanalysis. Genes contributing to four or more clinical cases with molecular diagnoses in the combined series are listed.

Table S1. Manual and semi-automated re-analysis results on the diagnoses from Cohort #1.

See the table at the end of this document.

Table S2. Reanalysis results in new diagnoses due to variant reclassification. *, these variants theoretically should have been included on the original report as VUSs. In the "Pathogenicity" column, "P" refers to pathogenic, "LP" refers to likely pathogenic. In the "Evidence level" column, "Nucleotide" refers to the evidence deriving from reports based on the identical variant; "Amino acid" refers to the evidence deriving from a different variant that results in the identical amino acid change; "Gene" refers to the evidence deriving from the gene itself, for example, establishment of new isoform or expansion of phenotypic association or associated inheritance pattern. In the "Comment" column, "atypical" refers to variant-specific atypical phenotypic presentations; "complex diagnosis" refers to complex patient phenotypes blurred by multilocus molecular diagnoses.

Variant on the original report	Gene of new diagnosis	HGVS nomenclature	Pathogenicity	Inheritance	New evidence source	Evidence level	Comment
Yes	*DEAF1*	NM_021008: c.676C>T (p.R226W)	LP	AR	Literature	Nucleotide	Multiple inheritance patterns
Yes	*STIM1*	NM_003156: c.910C>T (p.R304W)	P	AD	Literature	Nucleotide	Multiple inheritance patterns
Yes	*KCNJ2*	NM_000891: c.896A>G (p.E299G)	LP	AD	Literature	Nucleotide	-
Yes	*PNPO*	NM_018129: c.673C>T (p.R225C)	P	AR	Literature	Nucleotide	-
Yes	*HDAC8*	NM_018486: c.769C>T (p.P257S)	LP	XL	Literature	Nucleotide	-
Yes	*THOC6*	NM_024339: c.569G>A (p.G190E); [c.298T>A (p.W100R) +	LP; LP	AR	Internal	Nucleotide	-

		c.700G>C (p.V234L) + c.824G>A (p.G275D)]					
Yes	*L1CAM*	NM_000425: c.604G>A (p.D202N)	LP	XL	Literature	Nucleotide	-
Yes	*FHL1*	NM_001449: c.448T>C (p.C150R)	LP	AD	Literature	Amino acid	-
Yes	*COL1A2*	NM_000089: c.1342G>C (p.G448R)	LP	AD	Literature	Amino acid	-
Yes	*MYH7*	NM_000257: c.1141G>A (p.A381T)	LP	AD	Literature	Amino acid	-
Yes	*GABRG2*	NM_000816: c.316G>A (p.A106T)	LP	AD	ClinVar	Nucleotide	Potential complex diagnosis
Yes	*TBC1D24*	NM_001199107: c.457G>A (p.E153K)	LP	AR	Internal	Nucleotide	-
Yes	*RAB3GAP2*	NM_012414: c.1276C>T (p.R426C)	P	AR	Literature	Nucleotide	-
Yes	*ACTG2*	NM_001615: c.119G>A (p.R40H)	P	AR	Literature	Nucleotide	-
Yes	*SCN1A*	NM_001165963: c.4100A>T	LP	AD	Literature	Amino acid	-

		(p.N1367I)					
Yes	*RARS2*	NM_020320: c.419T>G (p.F140C); c.472_474del (p.158del)	P; P	AR	Internal/ ClinVar	Nucleotide	-
Yes	*KIF1A*	NM_004321: c.946C>T (p.R316W)	P	AD	Literature	Nucleotide	Complex diagnosis
Yes	*TPM1*	NM_001018005: c.688G>A (p.D230N)	P	AD	Literature	Nucleotide	-
Yes	*ABCA4*	NM_000350: c.6221G>T (p.G2074V); c.1804C>T (p.R602W)	LP; P	AR	Literature	Nucleotide	-
Yes	*RARS2*	NM_020320: c.419T>G (p.F140C); c.472_474del (p.158del)	P; P	AR	Internal/ Clinvar	Nucleotide	-
Yes	*FOXP1*	NM_032682: c.844_845del (p.V283fs)	P	AD	Literature/ Internal	Gene	Phenotypic expansion
Yes	*COQ2*	NM_015697: c.590G>A (p.R197H)	P	AR	Internal/ Literature	Amino acid	-
Yes	*KCNT1*	NM_020822: c.1193G>A	P	AD	Literature	Nucleotide	-

		(p.R398Q)					
No*	*CNGA3*	NM_001298: c.778G>A (p.D260N)	LP	AR	Literature	Nucleotide	-
No	*KMT2D*	NM_003482: c.15089G>A (p.R5030H)	LP	AD	Literature	Amino acid	Atypical/ Complex diagnosis
No	*ATP1A3*	NM_152296: c.2266C>T (p.R756C)	LP	AD	Literature	Amino acid	Atypical
No	*CBL*	NM_005188: c.1112A>C (p.Y371S)	LP	AD	Literature	Amino acid	Atypical
No	*ACVR1*	NM_001105: c.983G>A (p.G328E)	LP	AD	Literature	Nucleotide	Atypical
No	*DNM1L*	NM_012062: c.763_764dup (p.N256fs)	P	AR	Literature	Gene	Complex diagnosis/ Multiple inheritance patterns
No	*PEX6*	NM_000287: c.1802G>A (p.R601Q)	P	AR	Literature	Nucleotide	Atypical
No	*HMBS*	NM_000190: c.655G>T (p.A219S)	LP	AD	Literature	Amino acid	Complex diagnosis
No	*SCN8A*	NM_00133260: c.676C>G (p.R226G)	LP	AD	Literature	Gene	New isoform

No	*SCN8A*	NM_00133260: c.697G>A (p.V233I)	LP	AD	Literature	Gene	New isoform
No	*COL12A1*	NM_004370: c.7001T>C (p.I2334T)	LP	AD	Literature	Nucleotide	Potential complex diagnosis
No	*RAB27A*	NM_004580: c.244C>T (p.R82C)	LP	AR	Literature	Nucleotide	Atypical
No	*ATP1A3*	NM_152296: c.2839G>T (p.G947W)	LP	AD	Literature	Amino acid	Atypical
No	*TUBA1A*	NM_006009: c.167T>T (p.F56F)	P	AD	Literature	Nucleotide	Atypical
No	*DNM1L*	NM_012062: c.1207C>T (p.R403C)	P	AD	Internal/ Literature	Nucleotide	Multiple inheritance patterns

Table S3. Reanalysis results in new diagnoses due to updated clinical information.

Note that the case with the *ZNF335* variant was included in both cohorts.

Gene of new diagnosis (original diagnosis)	New phenotypes developed or provided	Diagnostic category	Key information
COL3A1	HP:0000978: Bruising susceptibility HP:0001058: Poor wound healing HP:0031157: Carotid cavernous fistula HP:0002579: Gastrointestinal dysmotility	Diagnosis	New phenotype
FLG	HP:0008064: Ichthyosis	Partial diagnosis	New phenotype
FLG (ANKRD11)	HP:0008064: Ichthyosis	Second diagnosis	New phenotype
SERPINA1	HP:0002086: Abnormality of the respiratory system HP:0002910: Elevated hepatic transaminases	Partial diagnosis	New phenotype
PHGDH	PGDH enzymatic deficiency (No HPO) HP:0012278: Abnormality of serine metabolism HP:0010895: Abnormality of glycine metabolism HP:0001511: Intrauterine growth retardation HP:0001263: Global developmental delay HP:0001276: Hypertonia HP:0100704: Cortical visual impairment	Diagnosis	Biochemical
PHGDH	PGDH enzymatic deficiency (No HPO) HP:0012278: Abnormality of serine metabolism HP:0001511: Intrauterine growth retardation HP:0001263: Global developmental delay HP:0001276: Hypertonia	Diagnosis	Biochemical
NDUFA4	Elevations in sebacic, suberic and 3-hydroxysebacic acids (No HPO)	Diagnosis	Biochemical
ZNF335	HP:0000252: Microcephaly HP:0007371: Corpus callosum atrophy	Diagnosis	New phenotype

	HP:0002188: Delayed CNS myelination		
KCNQ2	HP:0002066: Gait ataxia HP:0001250: Seizures HP:0002751: Kyphoscoliosis HP:0002188: Delayed CNS myelination HP:0012443: Abnormality of brain morphology	Diagnosis	New phenotype
DDC	HP:0004338: Abnormality of aromatic amino acid family metabolism HP:0003785: Decreased CSF homovanillic acid	Diagnosis	Biochemical
NGLY1 (RYR1)	N-glycanase 1 enzymatic deficiency (No HPO)	Second diagnosis	Biochemical
ACADS	Elevations in ethylmalonate, methylsuccinate, and butyrylcarnitine (No HPO)	Partial diagnosis	Biochemical
ACO2	Mitochondrial Aconitase 2 enzymatic deficiency (No HPO)	Diagnosis	Biochemical
RYR2	HP:0004756: Ventricular tachycardia	Diagnosis	New phenotype

Table S4. CNV diagnostic findings and risk factors identified by reanalysis. Three diagnostic CNVs from Cohort #2 were reported in the initial exome analysis and are therefore not included in the reanalysis counts in Figure 1.

CNV findings	Identified by cSNP or exome	Molecular diagnosis or risk factor	Identified in the initial analysis or reanalysis	Cohort #
PUF60 het deletion	cSNP	Diagnosis	Reanalysis	1
16p12.2 het deletion	cSNP	Risk factor	Reanalysis	1
22q11.21 het duplication	cSNP	Risk factor	Reanalysis	2
16p13.11 het deletion	cSNP	Risk factor	Reanalysis	2
TANGO2 hmz deletion	exome	Diagnosis	Reanalysis	2
TANGO2 hmz deletion	exome	Diagnosis	Reanalysis	2
PUF60 het deletion	cSNP	Diagnosis	Reanalysis	2
NDE1 het deletion + SNV	cSNP	Diagnosis	Initial	2
ITSN1 het deletion	cSNP	Diagnosis	Reanalysis	2
16p13.11 het deletion	cSNP	Risk factor	Reanalysis	2
SHANK3 het deletion	cSNP	Diagnosis	Reanalysis	2
CLDN1 hmz deletion	exome	Diagnosis	Reanalysis	2
Idic(15)	cSNP	Diagnosis	Reanalysis	2
STXBP1 het deletion	cSNP	Diagnosis	Reanalysis	2
Angelman/ Prader-Willi syndrome deletion	cSNP	Diagnosis	Initial	2
22q11.21 het duplication	cSNP	Risk factor	Reanalysis	2
TRIM37 hmz deletion	exome	Diagnosis	Reanalysis	2
15q11.2q13.1 het duplication	cSNP	Diagnosis	Reanalysis	2
Angelman/ Prader-Willi syndrome deletion	cSNP	Diagnosis	Initial	2
15q26.2q26.3 het deletion	cSNP	Diagnosis	Reanalysis	2
ARID1B het deletion	cSNP	Diagnosis	Reanalysis	2
15q13.1q13.3 het deletion (including *CHRNA7*)	cSNP	Diagnosis	Reanalysis	2
ABCA4 hmz deletion	exome	Diagnosis	Reanalysis	2
HNPP het deletion	cSNP	Risk factor	Reanalysis	2
17q12 het deletion (including *HNF1B*)	cSNP	Diagnosis	Reanalysis	2
1q42.2q43 het deletion	cSNP	Diagnosis	Reanalysis	2

Table S5. Reanalysis augmented diagnostic resolutions in cases with multilocus genomic variations. *, the condition associated with *TUBB4A* in this patient is leukodystrophy, hypomyelinating, 6 [MIM: 612438].

Gene for initial diagnosis	Gene for new diagnosis	Reason for new diagnosis	Relationship between new and previous diagnosis	Cohort #
RBM10	*SMARCA4*	New gene	Overlapping	1
	ZNF423	New gene	Overlapping	
ANKRD11	*FLG*	Clinical update	Distinct	1
FGFR3	*TLK2*	New gene	Distinct	1
-	*BAF1*	New gene	N/A	1
	SLC12A5	New gene	N/A	
ANKRD11	*AHDC1*	New gene	Overlapping	1
DOLK	*PURA*	New gene	Overlapping	2
PDE11A	*KMT2D*	Variant reclassification	Distinct	2
CACNA1A	*KMT2E*	New gene	Overlapping	2
VCL	*VARS2*	New gene	Distinct	2
FBN2	*ATP1A3*	Other	Distinct	2
NDE1	*FDXR*	New gene	Distinct	2
KMT2C	*TCF20*	New gene	Overlapping	2
NGLY1	*SLC12A5*	New gene	Overlapping	2
OPA1	*DNM1L*	Variant reclassification	Overlapping	2
FLG	*HMBS*	Variant reclassification	Distinct	2
FLG	*SCN8A*	Variant reclassification	Distinct	2
*TUBB4A**	*GRIN2D*	New gene	Overlapping	2
TRAPPC11	*DNAJC12*	New gene	Overlapping	2
ANKRD11	*HECW2*	New gene	Overlapping	2
DYSF	*KIF1A*	Variant reclassification	Distinct	2
KMT2A	*TCIRG1*	Family study	Distinct	2
EFTUD2	*GRIN2D*	New gene	Overlapping	2
16p11.2 deletion	*KAT6A*	New gene	Overlapping	2

Table S6. Previous molecular diagnoses overturned by reanalysis

Gene	Nomenclature	Previous category	Previous evidence	Re-classified category	New evidence
CRYGD	NM_006891:c.168C>G (p.Y56*)	Pathogenic	Literature, expected protein truncation	Likely Benign	Seen in ExAC 4 Hom, Seen in internal database 3 Hom
FBN1	NM_000138:c.3509G>A (p.R1170H)	Pathogenic	Literature	Variant of unknown significance	Seen in ExAC 142 Het; seen in internal database multiple times
DMD	NM_004006:c.4233+2C>T	Pathogenic	ESP5400 seen 3 times	Likely benign	PMID 26185613, 25163546, 23871722; Seen in ExAC 16 Hem.
LRP2	NM_004525:c.11092G>A (p.V3698M)	VUS (in trans with another variant contributing a probable diagnosis)	ESP5400 seen 63 times	Likely benign	Seen in ExAC 6 Hom
NF2	NM_000268:c.1786T>C (p.*596Q)	Pathogenic variant	Novel in control	Variant of unknown significance	Unaffected sibling is heterozygous for this change
POMGNT1	NM_017739:c.1298C>T (p.T433M)	VUS (in trans with another variant contributing a probable diagnosis)	ESP5400 seen 12 times	Likely benign	Seen in ExAC 4 Hom; seen in internal database 1 Hom (unaffected with this disease)

Table S7. Clinical impact of exome reanalysis. Below are the abbreviations used for column "Level of difficulty in communicating the updated exome report compared to the initial exome report.**"** Less: Less challenging because the parents were prepared for possible updates to the results and knew what to expect; More 1: More challenging in terms of establishing re-contact with the patient; More 2: More challenging because family does not seem to be aware that the updated report can be potentially important; More 3: More challenging in terms of explaining why the "negative" exome report issued before has become "positive" now; Similar: The level of difficulty is similar

Diagnostic findings	**Updated results communicated to the patient in a follow up appointment?**	**Time since updated results are issued to the provider (month)**	**Physician same?**	**Genetic counselor same?**	**Level of difficulty in communicating the updated exome report compared to the initial exome report**	**New medication started**	**Diet change instituted**	**Parents requested withdrawal of care**	**Modifications to existing treatment/management**	**Additional diagnostic procedures completed**	**Relatives had genetic testing for known familial mutation(s)**	**Information was used for reproductive planning such as**	**No change in medical management**
WWOX	Yes	12	Yes	N/A	Less						Yes	Yes	
ARID1B	Yes	64	Yes	Yes	More 1	Yes		Yes	Yes		Yes		

RBM10, SMARCA4, ZNF423	Yes	5	Yes	No	Less								Yes
WDR19	No. The patient was deceased.	62	Yes	Yes	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
COL3A1	Yes	60	Yes	No	Less	Yes				Yes			
KMT2A	Yes	25	Yes	Yes	Less					Yes			
ALPK3	Yes	5	No	No	More 2				Yes		Yes		
UDN2	Yes	2	Yes	Yes	Less								Yes
NKX2-1	Yes	55	Yes	N/A	Less	Yes			Yes	Yes			
C5orf42	Yes	59	Yes	N/A	Less								Yes
DVL3	No. Patient was deceased.	5	Yes	No	Less	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ASXL3	Yes	7	No	No	N/A								Yes
EPG5	No. Patient	5	Yes	Yes	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

	was deceased.												
NAA10	No	17	Yes	Yes	N/A								
ANKRD11, FLG	Yes	54	Yes	N/A	Less								Yes
DNM1L	Yes	6	Yes	Yes	More 1 + More 2					Yes			
TUBB4A	Yes	4	Yes	No	Similar				Yes				
PIK3R1	Yes	26	Yes	No	Less								Yes
HNRNPU	Yes	35	Yes	No	Similar					Yes			
OGT	Yes	4	Yes	No	Less								Yes
SLC6A1	No. Patient relocated.	3	Yes	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
NACC1	No. Lost to follow up. The parent was not interested	14	Yes	Yes	More 1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

	in discussing												
POLR1C	Yes	17	Yes	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SLC13A5	Yes	36	Yes	No	N/A		Yes		Yes	Yes	Yes		
DDX3X	Yes	23	Yes	N/A	More 3				Yes	Yes	Yes	Yes	
PUF60 deletion	No. Not yet.	10	Yes	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
KCNT1	No. Lost to follow up.	26	No	Yes	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SLC12A5,BRF1	Yes	8	Yes	No	Less								Yes
NR2F1	Yes	28	Yes	No	Less								Yes
KCNT1	Yes	44	Yes	N/A	More 3				Yes	Yes	Yes	Yes	
CDK13	Yes	1	Yes	No	More 2								Yes
KCNJ2	No. The patient relocated.	3	Yes	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
KCNT1	No. Lost to follow up. The	2	Yes	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

	parent was not interested in discussing												
KMT2A	Yes	45	Yes	No	Less								Yes
AHDC1, *ANKRD11*	Yes	34	Yes	No	Less								Yes
WDR45	No. The physician relocated after the results were called to the family and the family hasn't been seen.	49	No	Yes	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

PGAP1	Yes	23	Yes	Yes	Less				Yes		Yes		
WDR45	Yes	49	Yes	No	Less	Yes							
SOX5	Yes	26	Yes	No	Less								Yes
GMNN	Yes	19	Yes	Yes	Less				Yes				
HNF4A	No. The patient relocated.	4	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ZNF335	Yes	34	Yes	No	Less				Yes				

Table S8. Estimated cost of reanalysis. The laboratory-initiated reanalysis is calculated under a model of annual reanalysis over five years. The ×5 multiplier represents the *in silico* variant filtering and prioritization and variant triaging being repeated five times. The × 2.5 multiplier for "Report drafting" reflects that time investment for drafting an updated report for new diagnosis is on average 2.5 times of that used for drafting an original report.

	Physician initiated reanalysis - percent of time compared to initial analysis	Laboratory initiated reanalysis - percent of time compared to initial analysis
Sample receiving; DNA extraction; library preparation and next generation sequencing	0 %	0 %
Raw data processing (generation of bam and VCF)	0 %	0 % or depend on the new analysis algorithm, e.g. CNV analysis
In silico variant filtering and prioritization	100%	100% × 5 times
Clinical summary	100%	Minimum if phenotypic matching software used
Variant triaging	100%	1.8% × 5 times
Review of likely contributing variants, including Sanger studies	100%	13%
Report drafting	11.4% × 2.5	10.8% × 2.5
Counselor review and calling out	11.4%	10.8%

Table S1.

HGVS_g	Gene_name	HGVS_c	HGVS_p	Variant_type	Zygosity	Disease	Monoallelic_inheritance_mode	Biallelic_inheritance_mode	Novel_seen_in_population	Literature_report	Literature_report_allelic	pLI	Z_score_of_missense_intolerance	ExAC_Het_AC	ExAC_HMZ_AC	Patient_ID	Diagnosis_ID	Detected_by_reanalysis	Reason_for_nondetection_from_the_semi-automated_reanalysis_pipeline
NC_000023.10:g.41495833_41495834dupTT	*CASK*	NM_003688.3:c.913_914dupAA	NP_003679.2:p.(Gly306Argfs Ter5)	Frameshift	Het	FG syndrome 4 [MIM:300422]; Mental retardation and microcephaly with pontine and cerebellar hypoplasia [MIM:300749]	TRUE	FALSE	novel	FALSE	FALSE	0.9996752570	4.23993576	0	0	1	1	YES	.
NC_000016.9:g.78133783G>A	*WWOX*	NM_016373.3:c.107+1G>A	NA	Splice Donor Site	Het	Epileptic encephalopathy, early infantile, 28 [MIM:616211]; Spinocerebellar ataxia, autosomal recessive 12 [MIM:614322]	FALSE	TRUE	novel	FALSE	FALSE	0.0000000597	-3.503555708	0	0	2	2	YES	.
NC_000016.9:g.78420850G>A	*WWOX*	NM_016373.3:c.605+5G>A	NA	Splice Region	Het	Epileptic encephalopathy, early infantile, 28 [MIM:616211]; Spinocerebellar ataxia, autosomal recessive 12 [MIM:614322]	FALSE	TRUE	novel	FALSE	FALSE	0.0000000597	-3.503555708	0	0	2	2	YES	.
NC_000023.10:g.153296476delC	*MECP2*	NM_001110792.1:c.842delG	NP_001104262.1:p.(Gly281AlafsTer20)	Frameshift	Hem	Rett syndrome [MIM:312750]; Encephalopathy, neonatal severe [MIM:300673]; Mental retardation, X-linked syndromic, Lubs type [MIM:300260]; Mental retardation, X-linked, syndromic 13 [MIM:300055]	TRUE	TRUE	novel	TRUE	FALSE	0.6627975700	-0.921719735	0	0	3	3	YES	.
NC_000006.11:g.158567859G>A	*SERAC1*	NM_032861.3:c.442C>T	NP_116250.3:p.(Arg148Ter)	Substitution - Nonsense	Het	3-methylglutaconic aciduria with deafness, encephalopathy, and Leigh-like syndrome [MIM:614739]	FALSE	TRUE	seen	TRUE	FALSE	0.0044906110	0.712749102	3	0	4	4	YES	.
NC_000006.11:g.158567864delG	*SERAC1*	NM_032861.3:c.438delC	NP_116250.3:p.(Thr147Argfs Ter22)	Frameshift	Het	3-methylglutaconic aciduria with deafness, encephalopathy, and Leigh-like syndrome [MIM:614739]	FALSE	TRUE	novel	TRUE	FALSE	0.0044906110	0.712749102	0	0	4	4	YES	.
NC_000006.11:g.157510793C>T	*ARID1B*	NM_020732.3:c.3568C>T	NP_065783.3:p.(Gln1190Ter)	Substitution - Nonsense	Het	Mental retardation, autosomal dominant 12 [MIM:614562]; Coffin-Siris syndrome [MIM:135900]	TRUE	FALSE	novel	FALSE	FALSE	0.9996544970	3.58290219	0	0	5	5	YES	.
NC_000023.10:g.47038564T>C	*RBM10*	NM_005676.4:c.724+2T>C	NA	Splice Donor Site	Hem	TARP syndrome [MIM:311900]	TRUE	TRUE	novel	TRUE	FALSE	0.9997879430	4.903042691	0	0	6	6	YES	.
NC_000019.9:g.11132488G>A	*SMARCA4*	NM_001128844.1:c.2704G>A	NP_001122316.1:p.(Val902Met)	Substitution - Missense	Het	Mental retardation, autosomal dominant 16 [MIM:614609]; Coffin-Siris syndrome [MIM:135900]	TRUE	FALSE	novel	FALSE	FALSE	0.9999999790	8.576030568	0	0	6	7	YES	.
NC_000016.9:g.49670532C>T	*ZNF423*	NM_001271620.2:c.2351G>A	NP_001258549.1:p.(Gly784Glu)	Substitution - Missense	Het	Joubert syndrome 19 [MIM:614844]	TRUE	TRUE	novel	FALSE	FALSE	0.9961982970	3.186646396	0	0	6	8	YES	.
NC_000012.11:g.112915523A>G	*PTPN11*	NM_002834.3:c.922A>G	NP_002825.3:p.(Asn308Asp)	Substitution - Missense	Het	LEOPARD syndrome 1 [MIM:151100]; Metachondromatosis [MIM:156250]; Noonan syndrome 1 [MIM:163950]	TRUE	FALSE	seen	TRUE	FALSE	0.9997070420	3.424549666	1	0	7	9	YES	.
NC_000013.10:g.23915076_23915077delAT	*SACS*	NM_014363.5:c.2938_2939delAT	NP_055178.3:p.(Met980ValfsTer10)	Frameshift	Het	Spastic ataxia, Charlevoix-Saguenay type [MIM:270550]	FALSE	TRUE	novel	FALSE	FALSE	0.0000000052	-1.137692504	0	0	8	10	YES	.
NC_000013.10:g.23911581A>T	*SACS*	NM_014363.5:c.6434T>A	NP_055178.3:p.(Leu2145Ter)	Substitution - Nonsense	Het	Spastic ataxia, Charlevoix-Saguenay type [MIM:270550]	FALSE	TRUE	seen	FALSE	FALSE	0.0000000052	-1.137692504	2	0	8	10	YES	.
NC_000013.10:g.23906381_23906382delCT	*SACS*	NM_014363.5:c.11637_11638delAG	NP_055178.3:p.(Arg3879Serfs Ter12)	Frameshift	Hom	Spastic ataxia, Charlevoix-Saguenay type [MIM:270550]	FALSE	TRUE	novel	TRUE	FALSE	0.0000000052	-1.137692504	0	0	9	11	YES	.
NC_000004.11:g.39207247dupA	*WDR19*	NM_025132.3:c.781dupA	NP_079408.3:p.(Thr261Asnfs Ter13)	Frameshift	Het	Cranioectodermal dysplasia 4 [MIM:614378]; Short-rib thoracic dysplasia 5 with or without polydactyly [MIM:614376]; Nephronophthisis 13 [MIM:614377]; Senior-Loken syndrome 8 [MIM:616307]	FALSE	TRUE	seen	TRUE	FALSE	0.0000001510	0.024458152	2	0	10	12	YES	.
NC_000004.11:g.39274649G>A	*WDR19*	NM_025132.3:c.3533G>A	NP_079408.3:p.(Arg1178Gln)	Substitution - Missense	Het	Cranioectodermal dysplasia 4 [MIM:614378]; Short-rib thoracic dysplasia 5 with or without polydactyly [MIM:614376]; Nephronophthisis 13 [MIM:614377]; Senior-Loken syndrome 8 [MIM:616307]	FALSE	TRUE	seen	TRUE	FALSE	0.0000001510	0.024458152	9	0	10	12	YES	.
NC_000006.11:g.33409454_33409455delAG	*SYNGAP1*	NM_006772.2:c.2212_2213delAG	NP_006763.2:p.(Ser738Ter)	Substitution - Nonsense	Het	Mental retardation, autosomal dominant 5 [MIM:612621]	TRUE	FALSE	novel	TRUE	FALSE	0.9999933020	7.467233934	0	0	11	13	YES	.
NC_000003.11:g.25622065T>C	*RARB*	NM_000965.4:c.638T>C	NP_000956.2:p.(Leu213Pro)	Substitution - Missense	Het	Microphthalmia, syndromic 12 [MIM:615524]	TRUE	TRUE	novel	TRUE	FALSE	0.9986764470	3.140741149	0	0	12	14	YES	.
NC_000011.9:g.119148874A>T	*CBL*	NM_005188.3:c.1096-2A>T	NA	Splice Acceptor Site	Het	Juvenile myelomonocytic leukemia [MIM:607785]; Noonan syndrome-like disorder with or without juvenile myelomonocytic leukemia [MIM:613563]	TRUE	FALSE	novel	FALSE	FALSE	0.0098146350	1.638290017	0	0	13	15	NO	Novel variant without literature support or high impact pLI or missense Z score

NC_000002.11:g.189871135G>A	*COL3A1*	NM_000090.3:c.3158G>A	NP_000081.1:p.(Gly1053Asp)	Substitution - Missense	Het	Ehlers-Danlos syndrome, type IV [MIM:130050]	TRUE	TRUE	novel	TRUE	FALSE	0.9999999950	3.386434598	0	0	14	16	YES	.
NC_000001.10:g.152285081_152285084delACTG	*FLG*	NM_002016.1:c.2282_2285delCAGT	NP_002007.1:p.(Ser761CysfsTer36)	Frameshift	Het	Ichthyosis vulgaris [MIM:146700]	TRUE	TRUE	seen	TRUE	FALSE	NA	NA	1712	20	15	17	YES	.
NC_000011.9:g.67379696C>T	*NDUFV1*	NM_007103.3:c.1268C>T	NP_009034.2:p.(Thr423Met)	Substitution - Missense	Hom	Mitochondrial complex I deficiency [MIM:252010]	FALSE	TRUE	novel	TRUE	FALSE	0.0003000000	0.68117449	0	0	16	18	YES	.
NC_000010.10:g.135184082C>G	*ECHS1*	NM_004092.3:c.268G>C	NP_004083.3:p.(Gly90Arg)	Substitution - Missense	Het	Mitochondrial short-chain enoyl-CoA hydratase 1 deficiency [MIM:616277]	FALSE	TRUE	novel	FALSE	TRUE	0.8091200080	1.153658854	0	0	17	19	YES	.
NC_000010.10:g.135183412_135183413delTA	*ECHS1*	NM_004092.3:c.410_411delAT	NP_004083.3:p.(Tyr137CysfsTer7)	Frameshift	Het	Mitochondrial short-chain enoyl-CoA hydratase 1 deficiency [MIM:616277]	FALSE	TRUE	seen	FALSE	FALSE	0.8091200080	1.153658854	2	0	17	19	YES	.
NC_000003.11:g.33059974C>T	*GLB1*	NM_000404.2:c.1313G>A	NP_000395.2:p.(Gly438Glu)	Substitution - Missense	Het	GM1-gangliosidosis, type I [MIM:230500]; GM1-gangliosidosis, type II [MIM:230600]; GM1-gangliosidosis, type III [MIM:230650]; Mucopolysaccharidosis type IVB (Morquio) [MIM:253010]	FALSE	TRUE	novel	TRUE	FALSE	0.0000000941	0.56165896	0	0	18	20	YES	.
NC_000003.11:g.33114105C>T	*GLB1*	NM_000404.2:c.176G>A	NP_000395.2:p.(Arg59His)	Substitution - Missense	Het	GM1-gangliosidosis, type I [MIM:230500]; GM1-gangliosidosis, type II [MIM:230600]; GM1-gangliosidosis, type III [MIM:230650]; Mucopolysaccharidosis type IVB (Morquio) [MIM:253010]	FALSE	TRUE	seen	TRUE	FALSE	0.0000000941	0.56165896	5	0	18	20	YES	.
NC_000022.10:g.20049061G>A	*TANGO2*	NM_152906.6:c.460G>A	NP_690870.3:p.(Gly154Arg)	Substitution - Missense	Hom	Metabolic encephalomyopathic crises, recurrent, with rhabdomyolysis, cardiac arrhythmias, and neurodegeneration [MIM:616878]	FALSE	TRUE	seen	TRUE	FALSE	0.0005710000	0.27930112	19	0	19	21	YES	.
NC_000002.11:g.145161631_145161647delTAGCCCCGGTCGCAGTA	*ZEB2*	NM_014795.3:c.643_659delTACTGCGACCGGGGCTA	NP_055610.1:p.(Tyr215GlnfsTer18)	Frameshift	Het	Mowat-Wilson syndrome [MIM:235730]	TRUE	FALSE	novel	FALSE	FALSE	0.9995659540	4.945603428	0	0	20	22	YES	.
NC_000002.11:g.191125881C>T	*HIBCH*	NM_014362.3:c.517+1G>A	NA	Splice Donor Site	Het	3-hydroxyisobutryl-CoA hydrolase deficiency [MIM:250620]	FALSE	TRUE	novel	TRUE	FALSE	0.0000000117	-1.144661992	0	0	21	23	YES	.
NC_000002.11:g.191152340G>A	*HIBCH*	NM_014362.3:c.410C>T	NP_055177.2:p.(Ala137Val)	Substitution - Missense	Het	3-hydroxyisobutryl-CoA hydrolase deficiency [MIM:250620]	FALSE	TRUE	novel	TRUE	FALSE	0.0000000117	-1.144661992	0	0	21	23	YES	.
NC_000002.11:g.179300991G>A	*PRKRA*	NM_003690.4:c.665C>T	NP_003681.1:p.(Pro222Leu)	Substitution - Missense	Het	Dystonia 16 [MIM:612067]	FALSE	TRUE	seen	TRUE	FALSE	0.4196197320	2.047517525	18	0	22	24	NO	ExAC allele count higher than the stringent cut-off (5) used in the biallelic hypothesis
NC_000002.11:g.179301019A>G	*PRKRA*	NM_003690.4:c.637T>C	NP_003681.1:p.(Cys213Arg)	Substitution - Missense	Het	Dystonia 16 [MIM:612067]	FALSE	TRUE	novel	TRUE	FALSE	0.4196197320	2.047517525	0	0	22	24	NO	The pair from the same biallelc hypothesis not passing filter
NC_000023.10:g.53442118C>A	*SMC1A*	NM_006306.3:c.110G>T	NP_006297.2:p.(Gly37Val)	Substitution - Missense	Het	Cornelia de Lange syndrome 2 [MIM:300590]	TRUE	TRUE	novel	FALSE	FALSE	0.9999655730	6.678453682	0	0	23	25	YES	.
NC_000023.10:g.71681927G>A	*HDAC8*	NM_018486.2:c.932C>T	NP_060956.1:p.(Thr311Met)	Substitution - Missense	Het	Cornelia de Lange syndrome 5 [MIM:300882]	TRUE	TRUE	seen	TRUE	FALSE	0.8936060810	2.401122076	1	0	24	26	NO	Variant observed on time in ExAC
NC_000023.10:g.53265676C>T	*IQSEC2*	NM_001111125.2:c.3279G>A	NP_001104595.1:p.(Ser1093=)	Substitution - coding silent	Hem	Mental retardation, X-linked 1/78 [MIM:309530]	TRUE	TRUE	novel	FALSE	FALSE	0.9924260480	4.557178241	0	0	25	27	YES	.
NC_000008.10:g.38287269C>A	*FGFR1*	NM_023110.2:c.289G>T	NP_075598.2:p.(Gly97Cys)	Substitution - Missense	Het	Encephalocraniocutaneous lipomatosis [MIM:613001]; Hartsfield syndrome [MIM:615465]; Hypogonadotropic hypogonadism 2 with or without anosmia [MIM:147950]; Jackson-Weiss syndrome [MIM:123150]; Osteoglophonic dysplasia [MIM:166250]; Pfeiffer syndrome [MIM:101600]; Trigonocephaly 1 [MIM:190440]	TRUE	FALSE	novel	FALSE	TRUE	0.9673889830	2.733339236	0	0	26	28	YES	.
NC_000003.11:g.47889727C>T	*DHX30*	NM_138615.2:c.2344C>T	NP_619520.1:p.(Arg782Trp)	Substitution - Missense	Het	Neurodevelopmental disorder with severe motor impairment and absent language [MIM:617804]	TRUE	FALSE	novel	TRUE	FALSE	0.9999977560	6.740963489	0	0	27	29	YES	.
NC_000003.11:g.45561765A>T	*LARS2*	NM_015340.3:c.2269A>T	NP_056155.1:p.(Met757Leu)	Substitution - Missense	Het	Hydrops, lactic acidosis, and sideroblastic anemia [MIM:617021]; Perrault syndrome 4 [MIM:615300]	FALSE	TRUE	novel	FALSE	FALSE	0.0000019500	0.527755583	0	0	28	30	YES	.
NC_000003.11:g.45554650C>A	*LARS2*	NM_015340.3:c.1784C>A	NP_056155.1:p.(Ala595Asp)	Substitution - Missense	Het	Hydrops, lactic acidosis, and sideroblastic anemia [MIM:617021]; Perrault syndrome 4 [MIM:615300]	FALSE	TRUE	seen	FALSE	FALSE	0.0000019500	0.527755583	1	0	28	30	YES	.

NC_000011.9:g.118348811G>A	*KMT2A*	NM_001197104.1:c.3464G>A	NP_001184033.1:p.(Cys1155Tyr)	Substitution - Missense	Het	Wiedemann-Steiner syndrome [MIM:605130]	TRUE	FALSE	novel	TRUE	FALSE	1.0000000000	6.875216813	0	0	29	31	YES	.
NC_000016.9:g.56385308G>C	*GNAO1*	NM_020988.2:c.736G>C	NP_066268.1:p.(Glu246Gln)	Substitution - Missense	Het	Epileptic encephalopathy, early infantile, 17 [MIM:615473]; Neurodevelopmental disorder with involuntary movements [MIM:617493]	TRUE	FALSE	novel	FALSE	TRUE	0.9775973170	3.541536287	0	0	30	32	YES	.
NC_000015.9:g.85400439dupT	*ALPK3*	NM_020778.4:c.3076dupT	NP_065829.3:p.(Ser1026PhefsTer55)	Frameshift	Het	Severe Pediatric Cardiomyopathy [PMID:26846950]; Sporadic dilated cardiomyopathy. [PMID:28296976]	FALSE	TRUE	novel	FALSE	FALSE	0.0000000043	0.447993759	0	0	31	33	YES	.
NC_000015.9:g.85401144C>T	*ALPK3*	NM_020778.4:c.3781C>T	NP_065829.3:p.(Arg1261Ter)	Substitution - Nonsense	Het	Severe Pediatric Cardiomyopathy [PMID:26846950]; Sporadic dilated cardiomyopathy. [PMID:28296976]	FALSE	TRUE	seen	TRUE	FALSE	0.0000000043	0.447993759	5	0	31	33	YES	.
NC_000007.13:g.138960894G>A	*UBN2*	NM_173569.3:c.2024+1G>A	NA	Splice Donor Site	Het	Autism spectrum disorder [PMID:28263302]	TRUE	FALSE	novel	TRUE	FALSE	0.9991836130	-0.066606137	0	0	32	34	YES	.
NC_000014.8:g.51080061C>T	*ATL1*	NM_015915.4:c.715C>T	NP_056999.2:p.(Arg239Cys)	Substitution - Missense	Het	Neuropathy, hereditary sensory, type ID [MIM:613708]; Spastic paraplegia 3A, autosomal dominant [MIM:182600]	TRUE	FALSE	novel	TRUE	FALSE	0.9966688650	2.422367346	0	0	33	35	YES	.
NC_000005.9:g.167913505T>C	*RARS*	NM_002887.3:c.2T>C	NA	Start Codon	Het	Leukodystrophy, hypomyelinating, 9 [MIM:616140]	FALSE	TRUE	novel	FALSE	TRUE	0.0060351950	-0.582558779	0	0	34	36	YES	.
NC_000005.9:g.167943865G>A	*RARS*	NM_002887.3:c.1535G>A	NP_002878.2:p.(Arg512Gln)	Substitution - Missense	Het	Leukodystrophy, hypomyelinating, 9 [MIM:616140]	FALSE	TRUE	seen	TRUE	FALSE	0.0060351950	-0.582558779	2	0	34	36	YES	.
NC_000006.11:g.157488314_157488326delCGGCAGGTAACCT	*ARID1B*	NM_020732.3:c.3020_3025+7delCGGCAGGTAACCT	NA	Splice Donor Site	Het	Mental retardation, autosomal dominant 12 [MIM:614562]; Coffin-Siris syndrome [MIM:135900]	TRUE	FALSE	novel	FALSE	FALSE	0.9996544970	3.58290219	0	0	35	37	YES	.
NC_000014.8:g.36986817G>C	*NKX2-1*	NM_003317.3:c.782C>G	NP_003308.1:p.(Pro261Arg)	Substitution - Missense	Het	Chorea, hereditary benign [MIM:118700]; Choreoathetosis, hypothyroidism, and neonatal respiratory distress [MIM:610978]	TRUE	FALSE	novel	FALSE	TRUE	0.0680687690	3.397289945	0	0	36	38	YES	.
NC_000005.9:g.37177747T>C	*C5orf42*	NM_023073.3:c.5876A>G	NP_075561.3:p.(Glu1959Gly)	Substitution - Missense	Het	Joubert syndrome 17 [MIM:614615]; Orofaciodigital syndrome VI [MIM:277170]	FALSE	TRUE	novel	FALSE	FALSE	0.0000000000	-0.415550546	0	0	37	39	YES	.
NC_000005.9:g.37121860C>T	*C5orf42*	NM_023073.3:c.8882G>A	NP_075561.3:p.(Arg2961His)	Substitution - Missense	Het	Joubert syndrome 17 [MIM:614615]; Orofaciodigital syndrome VI [MIM:277170]	FALSE	TRUE	seen	FALSE	FALSE	0.0000000000	-0.415550546	3	0	37	39	YES	.
NC_000003.11:g.183887887delC	*DVL3*	NM_004423.3:c.1592delC	NP_004414.3:p.(Pro531LeufsTer137)	Frameshift	Het	Robinow syndrome, autosomal dominant 3 [MIM:616894]	TRUE	FALSE	novel	FALSE	FALSE	0.9893613670	3.612380387	0	0	38	40	YES	.
NC_000005.9:g.36962223A>G	*NIPBL*	NM_133433.3:c.459-2A>G	NA	Splice Acceptor Site	Het	Cornelia de Lange syndrome 1 [MIM:122470]	TRUE	FALSE	novel	TRUE	FALSE	1.0000000000	5.232105433	0	0	39	41	YES	.
NC_000002.11:g.176958362_176958365delGCCA	*HOXD13*	NM_000523.3:c.744_747delGCCA	NP_000514.2:p.(Gln248HisfsTer17)	Frameshift	Het	Brachydactyly-syndactyly syndrome [MIM:610713]; Brachydactyly, type D [MIM:113200]; Brachydactyly, type E [MIM:113300]; Syndactyly, type V [MIM:186300]; Synpolydactyly 1 [MIM:186000]	TRUE	FALSE	novel	FALSE	FALSE	0.2922712720	2.977617593	0	0	40	42	NO	Novel variant without literature support or high impact pLI or missense Z score
NC_000018.9:g.31319346_31319349delGACA	*ASXL3*	NM_030632.2:c.1978_1981delGACA	NP_085135.1:p.(Asp660Asnfs Ter16)	Frameshift	Het	Bainbridge-Ropers syndrome [MIM:615485]	TRUE	FALSE	novel	TRUE	FALSE	0.9999005850	-0.921705796	0	0	41	43	YES	.
NC_000008.10:g.61763035G>A	*CHD7*	NM_017780.3:c.5405-17G>A	NA	Splice Region	Het	CHARGE syndrome [MIM:214800]; Hypogonadotropic hypogonadism 5 with or without anosmia [MIM:612370]	TRUE	FALSE	novel	TRUE	FALSE	1.0000000000	2.242084967	0	0	42	44	YES	.
NC_000012.11:g.21958185_21958186insA	*ABCC9*	NM_020297.3:c.4512+746_4512+747insT	NA	Intronic	Het	Atrial fibrillation, familial, 12 [MIM:614050]; Cardiomyopathy, dilated, 1O [MIM:608569]; Hypertrichotic osteochondrodysplasia [MIM:239850]	TRUE	FALSE	novel	FALSE	FALSE	0.2066459010	4.970491185	0	0	43	45	YES	.
NC_000011.9:g.47463203C>T	*RAPSN*	NM_005055.4:c.872G>A	NP_005046.2:p.(Gly291Asp)	Substitution - Missense	Het	Fetal akinesia deformation sequence [MIM:208150]; Myasthenic syndrome, congenital, 11, associated with acetylcholine receptor deficiency [MIM:616326]	FALSE	TRUE	seen	TRUE	FALSE	0.1196149640	-0.393898376	3	0	43	46	YES	.
NC_000011.9:g.47469438C>T	*RAPSN*	NM_005055.4:c.457G>A	NP_005046.2:p.(Ala153Thr)	Substitution - Missense	Het	Fetal akinesia deformation sequence [MIM:208150]; Myasthenic syndrome, congenital, 11, associated with acetylcholine receptor deficiency [MIM:616326]	FALSE	TRUE	seen	FALSE	FALSE	0.1196149640	-0.393898376	2	0	43	46	YES	.
NC_000017.10:g.44248973delT	*KANSL1*	NM_001193466.1:c.540delA	NP_001180395.1:p.(Lys180AsnfsTer22)	Frameshift	Het	Koolen-De Vries syndrome [MIM:610443]	TRUE	FALSE	novel	TRUE	FALSE	0.9999652020	0.443929224	0	0	44	47	YES	.

NC_000003.11:g.4687362C>T	*ITPR1*	NM_001168272.1:c.805C>T	NP_001161744.1:p.(Arg269Trp)	Substitution - Missense	Het	Gillespie syndrome [MIM:206700]; Spinocerebellar ataxia 15 [MIM:606658]; Spinocerebellar ataxia 29, congenital nonprogressive [MIM:117360]	TRUE	TRUE	novel	TRUE	FALSE	1.0000000000	6.13165946	0	0	45	48	YES	.
NC_000002.11:g.241723197C>T	*KIF1A*	NM_001244008.1:c.757G>A	NP_001230937.1:p.(Glu253Lys)	Substitution - Missense	Het	Mental retardation, autosomal dominant 9 [MIM:614255]; Neuropathy, hereditary sensory, type IIC [MIM:614213]; Spastic paraplegia 30, autosomal recessive [MIM:610357]	TRUE	TRUE	novel	TRUE	FALSE	0.9999811750	5.419720448	0	0	46	49	YES	.
NC_000001.10:g.244218541G>T	*ZBTB18*	NM_205768.2:c.1465G>T	NP_991331.1:p.(Asp489Tyr)	Substitution - Missense	Het	Mental retardation, autosomal dominant 22 [MIM:612337]	TRUE	FALSE	novel	FALSE	FALSE	0.9617245740	3.641015982	0	0	47	50	YES	.
NC_000015.9:g.43900289G>A	*STRC*	NM_153700.2:c.3670C>T	NP_714544.1:p.(Arg1224Ter)	Substitution - Nonsense	Het	Deafness, autosomal recessive 16 [MIM:603720]	FALSE	TRUE	seen	FALSE	FALSE	0.0093358370	4.386660433	11	0	48	51	NO	Variant zygosity misattributed due to interference from repeat sequences, making one hit missing from a biallelic hypothesis
NC_000018.9:g.43531205C>T	*EPG5*	NM_020964.2:c.1253-1G>A	NA	Splice Acceptor Site	Het	Vici syndrome [MIM:242840]	FALSE	TRUE	novel	FALSE	FALSE	0.0000031500	-0.313944851	0	0	49	52	YES	.
NC_000018.9:g.43505706G>A	*EPG5*	NM_020964.2:c.2716C>T	NP_066015.2:p.(Gln906Ter)	Substitution - Nonsense	Het	Vici syndrome [MIM:242840]	FALSE	TRUE	seen	TRUE	FALSE	0.0000031500	-0.313944851	3	0	49	52	YES	.
NC_000023.10:g.153198002A>G	*NAA10*	NM_003491.3:c.215T>C	NP_003482.1:p.(Ile72Thr)	Substitution - Missense	Hem	Microphthalmia, syndromic 1 [MIM:309800]; Ogden syndrome [MIM:300855]	TRUE	TRUE	novel	TRUE	FALSE	0.1667480980	2.499825114	0	0	50	53	YES	.
NC_000016.9:g.89350552_89350555delCTTT	*ANKRD11*	NM_001256182.1:c.2398_2401delGAAA	NP_001243111.1:p.(Glu800AsnfsTer62)	Frameshift	Het	KBG syndrome [MIM:148050]	TRUE	FALSE	novel	TRUE	FALSE	0.9999999550	2.751549527	0	0	51	54	YES	.
NC_000001.10:g.152285081_152285084delACTG	*FLG*	NM_002016.1:c.2282_2285delCAGT	NP_002007.1:p.(Ser761CysfsTer36)	Frameshift	Het	Ichthyosis vulgaris [MIM:146700]	TRUE	TRUE	seen	TRUE	FALSE	NA	NA	1712	20	51	55	YES	.
NC_000022.10:g.41513200_41513203delCTCT	*EP300*	NM_001429.3:c.104_107delCTCT	NP_001420.2:p.(Ser35TyrfsTer12)	Frameshift	Het	Rubinstein-Taybi syndrome 2 [MIM:613684]	TRUE	FALSE	novel	TRUE	FALSE	1.0000000000	1.739038235	0	0	52	56	YES	.
NC_000014.8:g.51080061C>T	*ATL1*	NM_015915.4:c.715C>T	NP_056999.2:p.(Arg239Cys)	Substitution - Missense	Het	Neuropathy, hereditary sensory, type ID [MIM:613708]; Spastic paraplegia 3A, autosomal dominant [MIM:182600]	TRUE	FALSE	novel	TRUE	FALSE	0.9966688650	2.422367346	0	0	53	57	YES	.
NC_000022.10:g.24176330G>A	*SMARCB1*	NM_003073.3:c.1121G>A	NP_003064.2:p.(Arg374Gln)	Substitution - Missense	Het	Mental retardation, autosomal dominant 15 [MIM:614608]; Coffin-Siris syndrome [MIM:135900]	TRUE	FALSE	novel	TRUE	FALSE	0.9946363680	4.582978061	0	0	54	58	YES	.
NC_000012.11:g.32895600A>G	*DNM1L*	NM_012062.4:c.2072A>G	NP_036192.2:p.(Tyr691Cys)	Substitution - Missense	Het	Encephalopathy, lethal, due to defective mitochondrial peroxisomal fission 1 [MIM:614388]; Optic atrophy 5 [MIM:610708]	TRUE	TRUE	novel	FALSE	FALSE	0.2586120730	3.678334955	0	0	55	59	YES	.
NC_000017.10:g.72861044T>A	*FDXR*	NM_001258012.3:c.748A>T	NP_001244941.2:p.(Ile250Phe)	Substitution - Missense	Het	Auditory neuropathy and optic atrophy [MIM:617717]	FALSE	TRUE	novel	FALSE	FALSE	0.0004910000	0.702322133	0	0	56	60	YES	.
NC_000017.10:g.72862288C>T	*FDXR*	NM_001258012.3:c.601G>A	NP_001244941.2:p.(Val201Met)	Substitution - Missense	Het	Auditory neuropathy and optic atrophy [MIM:617717]	FALSE	TRUE	novel	FALSE	FALSE	0.0004910000	0.702322133	0	0	56	60	YES	.
NC_000023.10:g.149828138G>A	*MTM1*	NM_000252.2:c.1262G>A	NP_000243.1:p.(Arg421Gln)	Substitution - Missense	Hem	Myotubular myopathy, X-linked [MIM:310400]	TRUE	TRUE	novel	TRUE	FALSE	0.9987360080	2.231960099	0	0	57	61	YES	.
NC_000010.10:g.76784949_76784952delAACA	*KAT6B*	NM_012330.3:c.3606_3609delAACA	NP_036462.2:p.(Thr1203ArgfsTer21)	Frameshift	Het	Genitopatellar syndrome [MIM:606170]; SBBYSS syndrome [MIM:603736]	TRUE	FALSE	novel	TRUE	FALSE	0.9999999010	2.256273993	0	0	58	62	YES	.
NC_000014.8:g.92336713_92336714delCT	*FBLN5*	NM_006329.3:c.1201_1202delAG	NP_006320.2:p.(Ser401CysfsTer135)	Frameshift	Hom	Cutis laxa, autosomal dominant 2 [MIM:614434]; Cutis laxa, autosomal recessive, type IA [MIM:219100]; Macular degeneration, age-related, 3 [MIM:608895]	TRUE	TRUE	novel	FALSE	FALSE	0.9918595860	1.862027203	0	0	59	63	YES	.
NC_000023.10:g.76891451C>A	*ATRX*	NM_000489.4:c.4654G>T	NP_000480.3:p.(Val1552Phe)	Substitution - Missense	Hem	Alpha-thalassemia/mental retardation syndrome [MIM:301040]; Mental retardation-hypotonic facies syndrome, X-linked [MIM:309580]	TRUE	TRUE	novel	TRUE	FALSE	0.9999999690	3.189999668	0	0	60	64	YES	.
NC_000016.9:g.29825024dupC	*PRRT2*	NM_145239.2:c.649dupC	NP_660282.2:p.(Arg217ProfsTer8)	Frameshift	Het	Convulsions, familial infantile, with paroxysmal choreoathetosis [MIM:602066]; Episodic kinesigenic dyskinesia 1 [MIM:128200]; Seizures, benign familial infantile, 2 [MIM:605751]	TRUE	TRUE	novel	TRUE	FALSE	0.2660914910	0.656836164	0	0	61	65	YES	.

NC_000012.11:g.25398218G>C	*KRAS*	NM_004985.4:c.101C>G	NP_004976.2:p.(Pro34Arg)	Substitution - Missense	Het	Noonan syndrome 3 [MIM:609942]; Cardiofaciocutaneous syndrome [MIM:115150]; Leukemia, acute myelogenous	TRUE	FALSE	novel	TRUE	FALSE	0.0008000000	1.337421047	0	0	62	66	YES	.
NC_000012.11:g.52145307C>T	*SCN8A*	NM_014191.3:c.2300C>T	NP_055006.1:p.(Thr767Ile)	Substitution - Missense	Het	Cognitive impairment with or without cerebellar ataxia [MIM:614306]; Epileptic encephalopathy, early infantile, 13 [MIM:614558]; Seizures, benign familial infantile, 5 [MIM:617080]	TRUE	FALSE	novel	TRUE	FALSE	0.9999876360	7.857844875	0	0	63	67	YES	.
NC_000019.9:g.6495329A>G	*TUBB4A*	NM_006087.3:c.1181T>C	NP_006078.2:p.(Phe394Ser)	Substitution - Missense	Het	Dystonia 4, torsion, autosomal dominant [MIM:128101]; Leukodystrophy, hypomyelinating, 6 [MIM:612438]	TRUE	FALSE	novel	FALSE	TRUE	0.0136636790	6.012780757	0	0	64	68	YES	.
NC_000002.11:g.166201140G>A	*SCN2A*	NM_001040142.1:c.2638G>A	NP_001035232.1:p.(Ala880Thr)	Substitution - Missense	Het	Epileptic encephalopathy, early infantile, 11 [MIM:613721]; Seizures, benign familial infantile, 3 [MIM:607745]	TRUE	FALSE	novel	FALSE	FALSE	0.9999999310	6.88191849	0	0	65	69	YES	.
NC_000007.13:g.5569207G>C	*ACTB*	NM_001101.3:c.82C>G	NP_001092.1:p.(Arg28Gly)	Substitution - Missense	Het	Dystonia, juvenile-onset [MIM:607371]; Baraitser-Winter syndrome 1 [MIM:243310]	TRUE	FALSE	novel	FALSE	FALSE	0.9223365520	6.158905688	0	0	66	70	YES	.
NC_000005.9:g.67590398T>C	*PIK3R1*	NM_181523.2:c.1460T>C	NP_852664.1:p.(Phe487Ser)	Substitution - Missense	Het	Agammaglobulinemia 7, autosomal recessive [MIM:615214]; Immunodeficiency 36 [MIM:616005]; SHORT syndrome [MIM:269880]	TRUE	TRUE	novel	FALSE	FALSE	0.9938715040	2.635917096	0	0	67	71	YES	.
NC_000002.11:g.32366975G>A	*SPAST*	NM_014946.3:c.1496G>A	NP_055761.2:p.(Arg499His)	Substitution - Missense	Het	Spastic paraplegia 4, autosomal dominant [MIM:182601]	TRUE	FALSE	novel	TRUE	FALSE	0.9986175200	1.499469938	0	0	68	72	YES	.
NC_000001.10:g.245018807_245018808delAG	*HNRNPU*	NM_031844.2:c.2270_2271delCT	NP_114032.2:p.(Pro757ArgfsTer7)	Frameshift	Het	Epileptic encephalopathy, early infantile, 54 [MIM:617391]	TRUE	FALSE	novel	TRUE	FALSE	0.9997241620	4.178016314	0	0	69	73	YES	.
NC_000023.10:g.22112100G>C	*PHEX*	NM_000444.5:c.733-1G>C	NA	Splice Acceptor Site	Hem	Hypophosphatemic rickets, X-linked dominant [MIM:307800]	TRUE	TRUE	novel	TRUE	FALSE	0.9991062680	1.42837646	0	0	70	74	YES	.
NC_000003.11:g.48508395G>A	*TREX1*	NM_016381.4:c.506G>A	NP_057465.1:p.(Arg169His)	Substitution - Missense	Het	Aicardi-Goutieres syndrome 1, dominant and recessive [MIM:225750]; Chilblain lupus [MIM:610448]; Vasculopathy, retinal, with cerebral leukodystrophy [MIM:192315]	TRUE	TRUE	seen	TRUE	FALSE	0.0874812100	-1.37044239	19	0	70	75	YES	.
NC_000003.11:g.48508420_48508422dupGGC	*TREX1*	NM_016381.5:c.531_533dupGGC	NP_057465.1:p.(Ala178dup)	Insertion - In frame	Het	Aicardi-Goutieres syndrome 1, dominant and recessive [MIM:225750]; Chilblain lupus [MIM:610448]; Vasculopathy, retinal, with cerebral leukodystrophy [MIM:192315]	TRUE	TRUE	novel	TRUE	FALSE	0.0874812100	-1.37044239	0	0	70	75	YES	.
NC_000002.11:g.228566966dupA	*SLC19A3*	NM_025243.3:c.74dupT	NP_079519.1:p.(Ser26LeufsTer19)	Frameshift	Het	Thiamine metabolism dysfunction syndrome 2 (biotin- or thiamine-responsive encephalopathy type 2) [MIM:607483]	FALSE	TRUE	novel	TRUE	FALSE	0.0011170420	-0.941378515	0	0	71	76	YES	.
NC_000002.11:g.228566953_228566954dupTC	*SLC19A3*	NM_025243.3:c.81_82dupGA	NP_079519.1:p.(Met28ArgfsTer2)	Frameshift	Het	Thiamine metabolism dysfunction syndrome 2 (biotin- or thiamine-responsive encephalopathy type 2) [MIM:607483]	FALSE	TRUE	seen	TRUE	FALSE	0.0011170420	-0.941378515	2	0	71	76	YES	.
NC_000023.10:g.70787555C>T	*OGT*	NM_181673.2:c.2765C>T	NP_858059.1:p.(Thr922Ile)	Substitution - Missense	Hem	Mental retardation, X-linked 106 [MIM:300997]	TRUE	TRUE	novel	FALSE	FALSE	0.9998653150	5.736334677	0	0	72	77	YES	.
NC_000003.11:g.48612861C>T	*COL7A1*	NM_000094.3:c.6091G>A	NP_000085.1:p.(Gly2031Ser)	Substitution - Missense	Het	EBD inversa [MIM:226600]; EBD, Bart type [MIM:132000]; EBD, localisata variant; Epidermolysis bullosa dystrophica, AD [MIM:131750]; Epidermolysis bullosa pruriginosa [MIM:604129]; Epidermolysis bullosa, pretibial [MIM:131850]; Toenail dystrophy, isolated [MIM:607523]; Transient bullous of the newborn [MIM:131705]	TRUE	TRUE	seen	TRUE	FALSE	0.0000000000	1.739460961	1	0	73	78	YES	.
NC_000003.11:g.48609429C>T	*COL7A1*	NM_000094.3:c.7068+5G>A	NA	Splice Region	Het	EBD inversa [MIM:226600]; EBD, Bart type [MIM:132000]; EBD, localisata variant; Epidermolysis bullosa dystrophica, AD [MIM:131750]; Epidermolysis bullosa pruriginosa [MIM:604129]; Epidermolysis bullosa, pretibial [MIM:131850]; Toenail dystrophy, isolated [MIM:607523]; Transient bullous of the newborn [MIM:131705]	TRUE	TRUE	seen	TRUE	FALSE	0.0000000000	1.739460961	1	0	73	78	YES	.
NC_000003.11:g.11078563delG	*SLC6A1*	NM_003042.3:c.1711delG	NP_003033.3:p.(Val571SerfsTer46)	Frameshift	Het	Myoclonic-atonic epilepsy [MIM:616421]	TRUE	FALSE	novel	FALSE	FALSE	0.9985922260	4.494931039	0	0	74	79	YES	.

NC_000004.11:g.1801122C>T	*FGFR3*	NM_000142.4:c.251C>T	NP_000133.1:p.(Ser84Leu)	Substitution - Missense	Het	Achondroplasia [MIM:100800]; CATSHL syndrome [MIM:610474]; Crouzon syndrome with acanthosis nigricans [MIM:612247]; Hypochondroplasia [MIM:146000]; LADD syndrome [MIM:149730]; Muenke syndrome [MIM:602849]; SADDAN [MIM:616482]; Thanatophoric dysplasia, type I [MIM:187600]; Thanatophoric dysplasia, type II [MIM:187601]	TRUE	TRUE	novel	TRUE	FALSE	0.1472062870	1.36559782	0	0	75	80	YES	.
NC_000017.10:g.60689780C>T	*TLK2*	NM_006852.3:c.2107C>T	NP_006843.2:p.(Arg703Ter)	Substitution - Nonsense	Het	Intellectual disability [PMID:27479843]	TRUE	FALSE	novel	FALSE	FALSE	0.9999991050	5.792020175	0	0	75	81	YES	.
NC_000019.9:g.13246913C>T	*NACC1*	NM_052876.3:c.892C>T	NP_443108.1:p.(Arg298Trp)	Substitution - Missense	Het	Neurodevelopmental disorder with epilepsy, cataracts, feeding difficulties, and delayed brain myelination [MIM:617393]	TRUE	FALSE	novel	TRUE	FALSE	0.9550178420	5.387257845	0	0	76	82	YES	.
NC_000001.10:g.27094440G>T	*ARID1A*	NM_006015.4:c.3148G>T	NP_006006.3:p.(Asp1050Tyr)	Substitution - Missense	Het	Coffin-Siris syndrome 2 [MIM:614607]	TRUE	FALSE	novel	FALSE	FALSE	0.9999999740	4.614626098	0	0	77	83	YES	.
NC_000012.11:g.25378643C>T	*KRAS*	NM_004985.4:c.355G>A	NP_004976.2:p.(Asp119Asn)	Substitution - Missense	Het	Noonan syndrome 3 [MIM:609942]; Cardiofaciocutaneous syndrome [MIM:115150]; Leukemia, acute myelogenous	TRUE	FALSE	novel	FALSE	FALSE	0.0008000000	1.337421047	0	0	78	84	NO	Novel variant without literature support or high impact pLI or missense Z score
NC_000011.9:g.686986G>A	*DEAF1*	NM_021008.3:c.676C>T	NP_066288.2:p.(Arg226Trp)	Substitution - Missense	Hom	Mental retardation, autosomal dominant [MIM:615828], ?Dyskinesia, seizures, and intellectual developmental disorder [MIM:617171]	TRUE	TRUE	seen	TRUE	FALSE	0.0151446810	2.86731925	1	0	79	85	YES	.
NC_000023.10:g.13757139_13757142delGAAA	*OFD1*	NM_003611.2:c.400_403delGAAA	NP_003602.1:p.(Glu134IlefsTer10)	Frameshift	Het	Retinitis pigmentosa 23 [MIM:300424]; Joubert syndrome 10 [MIM:300804]; Orofaciodigital syndrome I [MIM:311200]; Simpson-Golabi-Behmel syndrome, type 2 [MIM:300209]	TRUE	TRUE	novel	TRUE	FALSE	0.9907657840	-0.22353337	0	0	80	86	YES	.
NC_000006.11:g.43488126delC	*POLR1C*	NM_203290.3:c.616delC	NP_976035.1:p.(Gln206LysfsTer48)	Frameshift	Het	Leukodystrophy, hypomyelinating, 11 [MIM:616494]; Treacher Collins syndrome 3 [MIM:248390]	TRUE	TRUE	novel	FALSE	FALSE	0.0263686980	-0.584517145	0	0	81	87	YES	.
NC_000006.11:g.43485062C>T	*POLR1C*	NM_203290.3:c.88C>T	NP_976035.1:p.(Pro30Ser)	Substitution - Missense	Het	Leukodystrophy, hypomyelinating, 11 [MIM:616494]; Treacher Collins syndrome 3 [MIM:248390]	TRUE	TRUE	novel	TRUE	FALSE	0.0263686980	-0.584517145	0	0	81	87	YES	.
NC_000023.10:g.76949325delT	*ATRX*	NM_000489.3:c.477delA	NP_000480.2:p.(Lys159AsnfsTer11)	Frameshift	Hom	Alpha-thalassemia/mental retardation syndrome [MIM:301040]; Mental retardation-hypotonic facies syndrome, X-linked [MIM:309580]	TRUE	TRUE	novel	FALSE	FALSE	0.9999999690	3.189999668	0	0	82	88	YES	.
NC_000017.10:g.6606350C>T	*SLC13A5*	NM_177550.4:c.655G>A	NP_808218.1:p.(Gly219Arg)	Substitution - Missense	Het	Epileptic encephalopathy, early infantile, 25 [MIM:615905]	FALSE	TRUE	seen	TRUE	FALSE	0.2365758560	0.926246496	26	0	83	89	YES	.
NC_000017.10:g.6590948A>G	*SLC13A5*	NM_177550.4:c.1475T>C	NP_808218.1:p.(Leu492Pro)	Substitution - Missense	Het	Epileptic encephalopathy, early infantile, 25 [MIM:615905]	FALSE	TRUE	novel	TRUE	FALSE	0.2365758560	0.926246496	0	0	83	89	YES	.
NC_000023.10:g.41204666_41204671dupCGTGAT	*DDX3X*	NM_001356.4:c.1180_1185dupCGTGAT	NP_001347.3:p.(Arg394_Asp395dup)	Insertion - In frame	Het	Mental retardation, X-linked 102 [MIM:300958]	TRUE	TRUE	novel	FALSE	TRUE	0.9979756450	5.259948572	0	0	84	90	YES	.
NC_000007.13:g.92132486dupA	*PEX1*	NM_000466.2:c.2097dupT	NP_000457.1:p.(Ile700TyrfsTer42)	Frameshift	Het	Heimler syndrome 1 [MIM:234580]; Peroxisome biogenesis disorder 1A (Zellweger) [MIM:214100]; Peroxisome biogenesis disorder 1B (NALD/IRD) [MIM:601539]	FALSE	TRUE	seen	TRUE	FALSE	0.0247064060	0.050098876	79	0	85	91	YES	.
NC_000007.13:g.92130876C>T	*PEX1*	NM_000466.2:c.2528G>A	NP_000457.1:p.(Gly843Asp)	Substitution - Missense	Het	Heimler syndrome 1 [MIM:234580]; Peroxisome biogenesis disorder 1A (Zellweger) [MIM:214100]; Peroxisome biogenesis disorder 1B (NALD/IRD) [MIM:601539]	FALSE	TRUE	seen	TRUE	FALSE	0.0247064060	0.050098876	33	0	85	91	YES	.
NC_000011.9:g.4095850C>T	*STIM1*	NM_003156.3:c.910C>T	NP_003147.2:p.(Arg304Trp)	Substitution - Missense	Het	Immunodeficiency 10 [MIM:612783]; Myopathy, tubular aggregate, 1 160565; Stormorken syndrome [MIM:185070]	TRUE	TRUE	novel	TRUE	FALSE	0.8618448680	2.113837185	0	0	86	92	YES	.
NC_000003.11:g.49027975dupC	*P4HTM*	NM_177938.2:c.286dupC	NP_808807.2:p.(Gln96ProfsTer29)	Frameshift	Het	.	TRUE	FALSE	novel	FALSE	FALSE	0.0664213780	2.304846461	0	0	87	93	NO	Gene lacking public evidence as a disease gene (the diagnosis is made by internal evidence)

NC_000003.11:g.49038916A>C	*P4HTM*	NM_177938.2:c.482A>C	NP_808807.2:p.(His161Pro)	Substitution - Missense	Het	.	TRUE	FALSE	novel	FALSE	FALSE	0.0664213780	2.304846461	0	0	87	93	NO	Gene lacking public evidence as a disease gene (the diagnosis is made by internal evidence)
NC_000022.10:g.51153476G>A	*SHANK3*	NM_001080420.1:c.2313+1G>A	NA	Splice Donor Site	Het	Phelan-McDermid syndrome [MIM:606232]	TRUE	FALSE	novel	FALSE	FALSE	0.9999159220	5.099912779	0	0	88	94	YES	.
NC_000009.11:g.138656907C>T	*KCNT1*	NM_020822.2:c.1066C>T	NP_065873.2:p.(Arg356Trp)	Substitution - Missense	Het	Epilepsy, nocturnal frontal lobe, 5 [MIM:615005]; Epileptic encephalopathy, early infantile, 14 [MIM:614959]	TRUE	FALSE	novel	FALSE	FALSE	0.8211831900	4.021017417	0	0	89	95	YES	.
NC_000016.9:g.3786772T>C	*CREBBP*	NM_004380.2:c.4439A>G	NP_004371.2:p.(Asp1480Gly)	Substitution - Missense	Het	Rubinstein-Taybi syndrome 1 [MIM:180849]	TRUE	FALSE	novel	FALSE	FALSE	1.0000000000	5.673235287	0	0	90	96	YES	.
NC_000014.8:g.105684059_105684061delTCT	*BRF1*	NM_001242786.1:c.1319_1321delAGA	NP_001229715.1:p.(Lys440del)	Deletion - In frame	Het	Cerebellofaciodental syndrome [MIM:616202]	FALSE	TRUE	novel	FALSE	FALSE	0.4215960560	0.620305911	0	0	91	97	YES	.
NC_000014.8:g.105695166A>G	*BRF1*	NM_001242786.1:c.434T>C	NP_001229715.1:p.(Leu145Pro)	Substitution - Missense	Het	Cerebellofaciodental syndrome [MIM:616202]	FALSE	TRUE	novel	FALSE	FALSE	0.4215960560	0.620305911	0	0	91	97	YES	.
NC_000020.10:g.44670074T>C	*SLC12A5*	NM_001134771.1:c.1030T>C	NP_001128243.1:p.(Phe344Leu)	Substitution - Missense	Het	Epileptic encephalopathy, early infantile, 34 [MIM:616645]; Epilepsy, idiopathic generalized, susceptibility to, 14 [MIM:616685]	TRUE	TRUE	novel	FALSE	FALSE	0.9999892440	5.585294126	0	0	91	98	YES	.
NC_000005.9:g.92921111T>C	*NR2F1*	NM_005654.5:c.382T>C	NP_005645.1:p.(Cys128Arg)	Substitution - Missense	Het	Bosch-Boonstra-Schaaf optic atrophy syndrome [MIM:615722]	TRUE	FALSE	novel	TRUE	FALSE	0.9631275340	6.064171468	0	0	92	99	YES	.
NC_000009.11:g.138651532G>A	*KCNT1*	NM_020822.2:c.862G>A	NP_065873.2:p.(Gly288Ser)	Substitution - Missense	Het	Epilepsy, nocturnal frontal lobe, 5 [MIM:615005]; Epileptic encephalopathy, early infantile, 14 [MIM:614959]	TRUE	FALSE	novel	TRUE	FALSE	0.8211831900	4.021017417	0	0	93	100	YES	.
NC_000007.13:g.40087446G>A	*CDK13*	NM_003718.4:c.2570G>A	NP_003709.3:p.(Gly857Glu)	Substitution - Missense	Het	Congenital heart defects, dysmorphic facial features, and intellectual developmental disorder [MIM:617360]	TRUE	FALSE	novel	FALSE	FALSE	0.9466801110	3.70743453	0	0	94	101	YES	.
NC_000010.10:g.89711899C>T	*PTEN*	NM_000314.4:c.517C>T	NP_000305.3:p.(Arg173Cys)	Substitution - Missense	Het	Bannayan-Riley-Ruvalcaba syndrome [MIM:153480]; Cowden syndrome 1 [MIM:158350]; Macrocephaly/autism syndrome [MIM:605309]; PTEN hamartoma tumor syndrome; VATER association with macrocephaly and ventriculomegaly [MIM:276950]	TRUE	TRUE	novel	TRUE	FALSE	0.9609649700	3.874522838	0	0	95	102	YES	.
NC_000017.10:g.68172076A>G	*KCNJ2*	NM_000891.2:c.896A>G	NP_000882.1:p.(Glu299Gly)	Substitution - Missense	Het	Andersen syndrome [MIM:170390]; Atrial fibrillation, familial, 9 [MIM:613980]; Short QT syndrome 3 [MIM:609622]	TRUE	FALSE	novel	FALSE	TRUE	0.9502583610	2.847231132	0	0	96	103	YES	.
NC_000016.9:g.30748648C>A	*SRCAP*	NM_006662.2:c.7287C>A	NP_006653.2:p.(Cys2429Ter)	Substitution - Nonsense	Het	Floating-Harbor syndrome [MIM:136140]	TRUE	FALSE	novel	FALSE	FALSE	0.9999999230	2.523837086	0	0	97	104	YES	.
NC_000009.11:g.138670657G>C	*KCNT1*	NM_020822.2:c.2718G>C	NP_065873.2:p.(Gln906His)	Substitution - Missense	Het	Epilepsy, nocturnal frontal lobe, 5 [MIM:615005]; Epileptic encephalopathy, early infantile, 14 [MIM:614959]	TRUE	FALSE	novel	FALSE	TRUE	0.8211831900	4.021017417	0	0	98	105	YES	.
NC_000011.9:g.118352809T>A	*KMT2A*	NM_001197104.1:c.4012+2T>A	NA	Splice Donor Site	Het	Wiedemann-Steiner syndrome [MIM:605130]	TRUE	FALSE	novel	FALSE	FALSE	1.0000000000	6.875216813	0	0	99	106	YES	.
NC_000001.10:g.43408994T>C	*SLC2A1*	NM_006516.2:c.19-2A>G	NA	Splice Acceptor Site	Het	Dystonia 9 [MIM:601042]; GLUT1 deficiency syndrome 1, infantile onset, severe [MIM:606777]; GLUT1 deficiency syndrome 2, childhood onset [MIM:612126]; Stomatin-deficient cryohydrocytosis with neurologic defects [MIM:608885]; Epilepsy, idiopathic generalized, susceptibility to, 12 [MIM:614847]	TRUE	TRUE	novel	TRUE	FALSE	0.9821003640	3.66072721	0	0	100	107	YES	.
NC_000023.10:g.13764524_13764529delGAGTAT	*OFD1*	NM_003611.2:c.604_609delGAGTAT	NP_003602.1:p.(Glu202_Tyr203del)	Deletion - In frame	Hem	Retinitis pigmentosa 23 [MIM:300424]; Joubert syndrome 10 [MIM:300804]; Orofaciodigital syndrome I [MIM:311200]; Simpson-Golabi-Behmel syndrome, type 2 [MIM:300209]	TRUE	TRUE	novel	TRUE	FALSE	0.9907657840	-0.22353337	0	0	101	108	YES	.
NC_000007.13:g.143048771C>T	*CLCN1*	NM_000083.2:c.2680C>T	NP_000074.2:p.(Arg894Ter)	Substitution - Nonsense	Het	Myotonia congenita, dominant [MIM:160800]; Myotonia congenita, recessive [MIM:255700]; Myotonia levior, recessive	TRUE	TRUE	seen	TRUE	FALSE	0.0000000000	0.106940753	381	2	102	109	YES	.
NC_000007.13:g.143047697G>A	*CLCN1*	NM_000083.2:c.2545G>A	NP_000074.2:p.(Ala849Thr)	Substitution - Missense	Het	Myotonia congenita, dominant [MIM:160800]; Myotonia congenita, recessive [MIM:255700]; Myotonia levior, recessive	TRUE	TRUE	seen	FALSE	TRUE	0.0000000000	0.106940753	46	1	102	109	YES	.

NC_000018.9:g.42531907G>A	*SETBP1*	NM_015559.2:c.2602G>A	NP_056374.2:p.(Asp868Asn)	Substitution - Missense	Het	Mental retardation, autosomal dominant 29 [MIM:616078]; Schinzel-Giedion midface retraction syndrome [MIM:269150]	TRUE	FALSE	novel	TRUE	FALSE	0.9975146790	2.051776107	0	0	102	110	YES	.
NC_000007.13:g.98527752G>A	*TRRAP*	NM_001244580.1:c.3316G>A	NP_001231509.1:p.(Glu1106Lys)	Substitution - Missense	Het	Autism spectrum disorder [PMID:27824329], Neurodevelopmental disorder [PMID:28628100]	TRUE	FALSE	novel	FALSE	FALSE	1.0000000000	10.16546125	0	0	103	111	YES	.
NC_000001.10:g.27876081delG	*AHDC1*	NM_001029882.3:c.2547delC	NP_001025053.1:p.(Ser850ProfsTer82)	Frameshift	Het	Neonatal hypotonia with sleep apnea, speech delay and intellectual disability [MIM:615829]	TRUE	FALSE	novel	TRUE	FALSE	0.9987570860	4.855128409	0	0	104	112	YES	.
NC_000016.9:g.89350944T>G	*ANKRD11*	NM_001256182.1:c.2006A>C	NP_001243111.1:p.(Asp669Ala)	Substitution - Missense	Het	KBG syndrome [MIM:148050]	TRUE	FALSE	novel	FALSE	FALSE	0.9999999550	2.751549527	0	0	104	113	YES	.
NC_000023.10:g.48933345_48933346delAT	*WDR45*	NM_007075.3:c.587_588delTA	NP_009006.2:p.(Ile196SerfsTer26)	Frameshift	Het	Neurodegeneration with brain iron accumulation 5 [MIM:300894]	TRUE	TRUE	novel	TRUE	FALSE	0.9556054060	1.802844525	0	0	105	114	YES	.
NC_000014.8:g.94844947C>T	*SERPINA1*	NM_001127701.1:c.1096G>A	NP_001121173.1:p.(Glu366Lys)	Substitution - Missense	Hom	Alpha-1-antitrypsin deficiency [MIM:613490]	FALSE	TRUE	seen	TRUE	FALSE	0.0000002130	-0.990296505	1499	13	106	115	YES	.
NC_000002.11:g.197708779_197708780insTA	*PGAP1*	NM_024989.3:c.2357_2358insTA	NP_079265.2:p.(Arg786SerfsTer35)	Frameshift	Het	Mental retardation, autosomal recessive 42 [MIM:615802]	FALSE	TRUE	novel	FALSE	FALSE	0.2394147940	0.171440484	0	0	107	116	YES	.
NC_000002.11:g.197757090A>G	*PGAP1*	NM_024989.3:c.1069T>C	NP_079265.2:p.(Trp357Arg)	Substitution - Missense	Het	Mental retardation, autosomal recessive 42 [MIM:615802]	FALSE	TRUE	novel	FALSE	FALSE	0.2394147940	0.171440484	0	0	107	116	YES	.
NC_000023.10:g.48934185T>G	*WDR45*	NM_007075.3:c.345-2A>C	NA	Splice Acceptor Site	Het	Neurodegeneration with brain iron accumulation 5 [MIM:300894]	TRUE	TRUE	novel	TRUE	FALSE	0.9556054060	1.802844525	0	0	108	117	YES	.
NC_000016.9:g.81398716G>A	*GAN*	NM_022041.3:c.1373+1G>A	NA	Splice Donor Site	Het	Giant axonal neuropathy-1 [MIM:256850]	FALSE	TRUE	novel	FALSE	FALSE	0.0828172350	-0.195922158	0	0	109	118	YES	.
NC_000016.9:g.81398610T>C	*GAN*	NM_022041.3:c.1268T>C	NP_071324.1:p.(Ile423Thr)	Substitution - Missense	Het	Giant axonal neuropathy-1 [MIM:256850]	FALSE	TRUE	novel	TRUE	FALSE	0.0828172350	-0.195922158	0	0	109	118	YES	.
NC_000022.10:g.41573207G>C	*EP300*	NM_001429.3:c.5492G>C	NP_001420.2:p.(Arg1831Thr)	Substitution - Missense	Het	Rubinstein-Taybi syndrome 2 [MIM:613684]	TRUE	FALSE	novel	FALSE	FALSE	1.0000000000	1.739038235	0	0	110	119	YES	.
NC_000017.10:g.46024035C>T	*PNPO*	NM_018129.3:c.673C>T	NP_060599.1:p.(Arg225Cys)	Substitution - Missense	Hom	Pyridoxamine 5'-phosphate oxidase deficiency [MIM:610090]	FALSE	TRUE	seen	TRUE	FALSE	0.0000207000	1.441893949	3	0	111	120	YES	.
NC_000012.11:g.23757425C>A	*SOX5*	NM_152989.4:c.1021G>T	NP_694534.1:p.(Gly341Ter)	Substitution - Nonsense	Het	Lamb-Shaffer syndrome [MIM:616803]	TRUE	FALSE	novel	TRUE	FALSE	0.9982976880	3.286692964	0	0	112	121	YES	.
NC_000021.8:g.38877745C>T	*DYRK1A*	NM_101395.2:c.1399C>T	NP_567824.1:p.(Arg467Ter)	Substitution - Nonsense	Het	Mental retardation, autosomal dominant 7 [MIM:614104]	TRUE	FALSE	novel	TRUE	FALSE	0.9989898680	3.625907038	0	0	113	122	YES	.
NC_000010.10:g.120905755delA	*SFXN4*	NM_213649.1:c.930delT	NP_998814.1:p.(Ile310MetfsTer33)	Frameshift	Hom	Combined oxidative phosphorylation deficiency 18 [MIM:615578]	FALSE	TRUE	novel	FALSE	FALSE	0.0003650000	1.005478948	0	0	114	123	YES	.
NC_000001.10:g.120284440G>A	*PHGDH*	NM_006623.3:c.1129G>A	NP_006614.2:p.(Gly377Ser)	Substitution - Missense	Het	Neu-Laxova syndrome 1 [MIM:256520]; Phosphoglycerate dehydrogenase deficiency [MIM:601815]	FALSE	TRUE	novel	TRUE	FALSE	0.0013383630	0.498654204	0	0	115	124	NO	Only one of the two alleles from a biallelic hypothesis detected
NC_000006.11:g.24777490A>T	*GMNN*	NM_015895.4:c.16A>T	NP_056979.1:p.(Lys6Ter)	Substitution - Nonsense	Het	Meier-Gorlin syndrome 6 [MIM:616835]	TRUE	FALSE	novel	FALSE	FALSE	0.5676373790	-0.082996592	0	0	116	125	NO	Novel variant without literature support or high impact pLI or missense Z score
NC_000020.10:g.43034835C>T	*HNF4A*	NM_000457.4:c.253C>T	NP_000448.3:p.(Arg85Trp)	Substitution - Missense	Het	Fanconi renotubular syndrome 4, with maturity-onset diabetes of the young [MIM:616026]; MODY, type I [MIM:125850]	TRUE	FALSE	novel	TRUE	FALSE	0.9722388360	1.737104671	0	0	117	126	YES	.
NC_000020.10:g.44578004C>A	*ZNF335*	NM_022095.3:c.3787G>T	NP_071378.1:p.(Glu1263Ter)	Substitution - Nonsense	Het	Microcephaly 10, primary, autosomal recessive [MIM:615095]	FALSE	TRUE	novel	TRUE	FALSE	0.0063880720	-0.059156639	0	0	118	127	YES	.
NC_000020.10:g.44581308_44581311delTCAC	*ZNF335*	NM_022095.3:c.2744_2747delGTGA	NP_071378.1:p.(Ser915ThrfsTer3)	Frameshift	Het	Microcephaly 10, primary, autosomal recessive [MIM:615095]	FALSE	TRUE	seen	TRUE	FALSE	0.0063880720	-0.059156639	13	0	118	127	YES	.